PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C07H 21/04, C12N 5/00, 15/00
C12P 7/64, 21/04

A1

(11) International Publication Number: WO 94/10189

(43) International Publication Date: 11 May 1994 (11.05.94)

(21) International Application Number: PCT/US93/10526

(22) International Filing Date: 2 November 1993 (02.11.93)

(30) Priority data:
07/971,182 2 November 1992 (02.11.92) US

(71) Applicant: CALGENE, INC. [US/US]; 1920 Fifth Street, Davis, CA 95616 (US).

(72) Inventors: KNAUF, Vic., C. ; 1013 Hillview Lane, Winters, CA 95694 (US). THOMPSON, Gregory, A. ; 5127 Cowell Boulevard, Davis, CA 95616 (US).

(74) Agents: LASSEN, Elizabeth et al.; Calgene, Inc., 1920 Fifth Street, Davis, CA 95616 (US).

(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: PLANT FATTY ACID SYNTHASES

(57) Abstract

By this invention, compositions and methods of use related to β-ketoacyl-ACP synthase, hereinafter also referred to as "synthase", are provided. Also of interest are methods and compositions of amino acid and nucleic acid sequences related to biologically active plant synthase(s) factors. Amino acid and nucleic acid for synthase protein factors are provided, as well as methods to utilize such sequences in constructs for production of genetically engineered plants having altered fatty acid compositions. In addition, uses of non-plant synthase proteins in plant genetic engineering methods are also considered.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

PLANT FATTY ACID SYNTHASES

This application is a continuation-in-part of USSN 07/971,182 filed on November 2, 1992.

Field of Invention

The present invention is directed to synthase enzymes relevant to fatty acid synthesis, amino acid and nucleic acid sequences related thereto, and methods of using such compositions in plants.

Introduction

Background

Plant oils are used in a variety of industrial and edible uses. Novel vegetable oils compositions and/or improved means to obtain oils compositions, from biosynthetic or natural plant sources, are needed. Depending upon the intended oil use, various different fatty acid compositions are desired.

For example, in some instances having an oilseed with a higher ratio of oil to seed meal would be useful to obtain a desired oil at lower cost. This would be typical of a high value oil product. In some instances, having an oilseed with a lower ratio of oil to seed meal would be useful to lower caloric content. In other uses, edible plant oils with a higher percentage of unsaturated fatty acids are desired for cardio-vascular health reasons. And alternatively, temperate substitutes for high saturate tropical oils such as palm and coconut, would also find uses in a variety of industrial and food applications.

One means postulated to obtain such oils and/or modified fatty acid compositions is through the genetic engineering of plants. However, in order to genetically engineer plants one must have in place the means to transfer genetic material to the plant in a stable and heritable manner. Additionally, one must have nucleic acid sequences capable of producing the desired phenotypic result, regulatory regions capable of directing the correct

application of such sequences, and the like. Moreover, it should be appreciated that in order to produce a desired phenotype requires that the Fatty Acid Synthetase (FAS) pathway of the plant is modified to the extent that the ratios of reactants are modulated or changed.

Higher plants appear to synthesize fatty acids via a common metabolic pathway. In developing seeds, where fatty acids attached to triglycerides are stored as a source of energy for further germination, the FAS pathway is located in the proplastids. The first step is the formation of acetyl-ACP (acyl carrier protein) from acety-CoA and ACP catalyzed by the enzyme, acetyl-CoA:ACP transacylase (ATA). Elongation of acetyl-ACP to 16- and 18- carbon fatty acids involves the cyclical action of the following sequence of reactions: condensation with a two-carbon unit from malonyl-ACP to form a ß-ketoacyl-ACP (ß-ketoayl-ACP synthase), reduction of the keto-function to an alcohol (ß-ketoacyl-ACP reductase), dehydration to form an enoyl-ACP (ß-hydroxyacyl-ACP dehydrase), and finally reduction of the enoyl-ACP to form the elongated saturated acyl-ACP (enoyl-ACP reductase). ß-ketoacyl-ACP synthase I, catalyzes elongation up to palmitoyl-ACP (C16:0), whereas ß-ketoacyl-ACP synthase II catalyzes the final elongation to stearoyl-ACP (C18:0). Common plant unsaturated fatty acids, such as oleic, linoleic and α-linolenic acids found in storage triglycerides, originate from the desaturation of stearoyl-ACP to form oleoyl-ACP (C18:1) in a reaction catalyzed by a soluble plastid Δ-9 desaturase (also often referred to as "stearoyl-ACP desaturase"). Molecular oxygen is required for desaturation in which reduced ferredoxin serves as an electron co-donor. Additional desaturation is effected sequentially by the actions of membrane bound Δ-12 desaturase and Δ-15 desaturase. These "desaturases" thus create mono- or polyunsaturated fatty acids respectively.

A third ß-ketoacyl-ACP synthase has been reported in *S. oleracea* leaves having activity specific toward very short acyl-ACPs. This acetoacyl-ACP synthase or "ß-ketoacyl-ACP" synthase III has a preference to acetyl-CoA over acetyl-ACP, Jaworski, J.G., *et al.*, *Plant Phys.* (1989)

90:41-44. It has been postulated that this enzyme may be an alternate pathway to begin FAS, instead of ATA.

Obtaining nucleic acid sequences capable of producing a phenotypic result in FAS, desaturation and/or incorporation of fatty acids into a glycerol backbone to produce an oil is subject to various obstacles including but not limited to the identification of metabolic factors of interest, choice and characterization of a protein source with useful kinetic properties, purification of the protein of interest to a level which will allow for its amino acid sequencing, utilizing amino acid sequence data to obtain a nucleic acid sequence capable of use as a probe to retrieve the desired DNA sequence, and the preparation of constructs, transformation and analysis of the resulting plants.

Thus, the identification of enzyme targets and useful plant sources for nucleic acid sequences of such enzyme targets capable of modifying fatty acid compositions are needed. Ideally an enzyme target will be amenable to one or more applications alone or in combination with other nucleic acid sequences, relating to increased/decreased oil production, the ratio of saturated to unsaturated fatty acids in the fatty acid pool, and/or to novel oils compositions as a result of the modifications to the fatty acid pool. Once enzyme target(s) are identified and qualified, quantities of protein and purification protocols are needed for sequencing. Ultimately, useful nucleic acid constructs having the necessary elements to provide a phenotypic modification and plants containing such constructs are needed.

Brief Description of the Figures

Figure 1 provides cDNA and translated amino acid sequences of a 50kD *R. communis* synthase factor B gene. Preliminary cDNA sequence and the corresponding translational peptide sequence derived from the cDNA clone, pCGN2765 (2-8), which encodes the 50kD synthase protein is shown. The cDNA includes both the postulated transit

peptide sequence (amino acids 1-42) and the sequence encoding the mature protein.

Figure 2 provides the *R.communis* synthase factor B 2-8 sequence with additional 3' untranslated sequence.

Figure 3 provides cDNA and translated amino acid sequences of a *R. communis* 46kD synthase factor A gene.

Figures 4 and 5 provide cDNA and translated amino acid sequences of *Brassica* synthase factor B genes.

Figure 4 provides sequences of the cDNA insert of pCGN3248.

Figure 5 provides sequences of clone 4A.

Figure 6 provides cDNA sequence of a *Brassica* synthase factor A gene. Comparison of the translated amino acid sequence to the *R. communis* factor A sequence reveals a possible frame shift mutation in the region near nucleotide 1120.

Figure 7 provides translated amino acid sequence of nucleotides 79-1119 of the *Brassica* synthase A gene sequence shown in Fig. 6.

Figure 8 provides translated amino acid sequence of nucleotides 1127-1606 of the *Brassica* synthase A gene sequence shown in Fig. 6.

Figure 9 provides approximately 2 kb of genomic sequence of Bce4.

Figure 10 provides a cDNA sequence and the corresponding translational peptide sequence derived from *C. tinctorius* desaturase. The cDNA includes both the plastid transit peptide sequence (amino acids 1-33) and the sequence encoding the mature protein.

Summary of the Invention

By this invention, compositions and methods of use related to ß-ketoacyl-ACP synthase, hereinafter also referred to as "synthase", are provided. Also of interest are methods and compositions of amino acid and nucleic acid sequences related to biologically active plant synthase(s). Various plant synthase factor A and B proteins are described in WO 92/03564 which is hereby incorporated by reference in its entirety. As is described herein,

synthase III constructs for expression in plant cells may also be used, either alone, or in conjunction with other plant synthase or fatty acid biosynthesis gene sequences to provide enhanced oil yields and/or altered compositions of plant seed oil.

Nucleic acid sequences encoding a synthase protein required for synthase activity in a host cell may be employed in nucleic acid constructs to modulate the amount of synthase activity present in the host cell. A synthase may be produced in host cells for harvest or as a means of effecting a contact between the synthase and its substrate. Host cells include prokaryotes and/or eukaryotes. Plant host cells containing recombinant constructs encoding a synthase protein, as well as plants and cells containing modified levels of synthase protein(s) are also provided. Additional nucleic acid sequences, such as those encoding transit peptides, may also be used, particularly where a full length close for a particular synthase protein is not available or where a non-plant synthase sequence is used.

In addition, nucleic acid constructs may be designed to decrease expression of endogenous synthase in a plant cell as well. One example is the use of an anti-sense synthase sequence under the control of a promoter capable of expression in at least those plant cells which normally produce the enzyme.

Additionally, one may wish to coordinate expression of a synthase protein with the expression of other synthase proteins or other introduced sequences encoding enzymes related to fatty acid synthesis. For example, coordinated expression of synthase factor A and synthase factor B may be desirable to provide optimal synthase-II type activity in plant cells. Furthermore, coordinated expression of the synthase III gene with plant synthase proteins may also be desired. Examples of other enzymes related to fatty acid synthesis which may find use in conjunction with synthase proteins include plant thioesterases, especially medium-chain thioesterases, desaturases, especially Δ-9 desaturases, and the like. When nucleic acid constructs encoding such factors are prepared for introduction into a

plant cell, the transcriptional initiation regions may be different from each other.

Furthermore, uses of non-plant synthase protein sequences in plant cells are considered herein. Such sequences may be used alone or in conjunction with plant synthase proteins. For example, constructs for expression of an *E. coli* synthase III protein in plant cells are described. Such constructs may be modified to provide optimal codons for expression in plant cells, as well as to provide transit peptide sequences to target the synthase protein to plastids for effect on the plant fatty acid synthesis reactions.

Detailed Description of the Invention

A plant synthase of this invention includes any sequence of amino acids, polypeptide, peptide fragment or other protein preparation, whether derived in whole or in part from natural or synthetic sources which demonstrates the ability to catalyze a condensation reaction between an acyl-ACP or acyl-CoA having a chain length of C_2 to C_{16} and malonyl-ACP in a plant host cell. A plant synthase will be capable of catalyzing a synthase reaction in a plant host cell, i.e., *in vivo*, or in a plant cell-like environment, i.e., *in vitro*. Typically, a plant synthase will be derived in whole or in part from a natural plant source.

In addition, synthase from other sources such as bacteria or lower plants, may also be useful in plants and thus be considered a plant synthase in this invention. For example, the *E. coli* synthase protein encoded by the *fab*B gene is shown herein to have homology to plant synthase proteins. In *E. coli*, synthase I enzymatic activity is provided by a homodimer of the *fab*B gene product. Of particular interest is a gene for *E. coli* synthase III (*fab*H). Constructs for expression of the bacterial gene in plant cells will include fusion constructs to incorporate chloroplast transit peptide sequences, such that the *E. coli* synthase III gene product is directed to the site of fatty acid synthesis. In this manner, the overall lipid

yield may be increased by enhancing the first step in the FAS pathway.

Synthase I demonstrates preferential activity towards acyl-ACPs having shorter carbon chains, C_2-C_{14}; synthase II demonstrates preferential activity towards acyl-ACPs having longer carbon chains, C_{14}-C_{16}. Synthase III demonstrates preferential activity towards acyl-CoAs having very short carbon chains, C_2 to C_6. Other plant synthases may also find applicability by this invention, including synthase III type activities.

Synthases include modified amino acid sequences, such as sequences which have been mutated, truncated, increased and the like, as well as such sequences which are partially or wholly artificially synthesized. Synthases and nucleic acid sequences encoding synthases may be obtained by partial or homogenous purification of plant extracts, protein modeling, nucleic acid probes, antibody preparations, or sequence comparisons, for example. Once purified synthase is obtained, it may be used to obtain other plant synthases by contacting an antibody specific to *R. communis* synthase with a plant synthase under conditions conducive to the formation of an antigen:antibody immunocomplex and the recovery of plant synthase which reacts thereto. Once the nucleic acid sequence encoding a synthase is obtained, it may be employed in probes for further screening or used in genetic engineering constructs for transcription or transcription and translation in host cells, especially plant host cells.

Recombinant constructs containing a nucleic acid sequence encoding a synthase and a heterologous nucleic acid sequence of interest may be prepared. By heterologous is meant any sequence which is not naturally found joined to the synthase sequence. Hence, by definition, a sequence joined to any modified synthase is not a wild-type sequence. Other examples include a synthase from one plant source which is integrated into the genome of a different plant host.

Constructs may be designed to produce synthase in either prokaryotic or eukaryotic cells. The increased

expression of a synthase in a plant cell or decreased amount of endogenous synthase observed in a plant cell are of special interest. Moreover, in a nucleic acid construct for integration into a plant host genome, the synthase may be found in a "sense" or "anti-sense" orientation in relation to the direction of transcription. Thus, nucleic acids may encode biologically active synthases or sequences complementary to the sequence encoding a synthase to inhibit the production of endogenous plant synthase. By transcribing and translating a sense sequence in a plant host cell, the amount of synthase available to the plant FAS complex is increased. By transcribing or transcribing and translating an anti-sense sequence in a plant host cell, the amount of the synthase available to the plant FAS is decreased. Ideally, the anti-sense sequence is very highly homologous to the endogenous sequence. Other manners of decreasing the amount of synthase available to FAS may be employed, such as ribozymes or the screening of plant cells transformed with constructs containing sense sequences which in fact act to decrease synthase expression, within the scope of this invention. Other analogous methods may be applied by those of ordinary skill in the art.

Synthases may be used, alone or in combination, to catalyze the elongating condensation reactions of fatty acid synthesis depending upon the desired result. For example, rate influencing synthase activity may reside in synthase I-type, synthase II-type, synthase III-type or in a combination of these enzymes. Furthermore, synthase activities may rely on a combination of the various synthase factors as described in WO 92/03564.

Constructs which contain elements to provide the transcription and translation of a nucleic acid sequence of interest in a host cell are "expression cassettes". Depending upon the host, the regulatory regions will vary, including regions from structural genes from viruses, plasmid or chromosomal genes, or the like. For expression in prokaryotic or eukaryotic microorganisms, particularly unicellular hosts, a wide variety of constitutive or

regulatable promoters may be employed. Among transcriptional initiation regions which have been described are regions from bacterial and yeast hosts, such as *E. coli, B. subtilis, Saccharomyces cerevisiae*, including genes such as ß-galactosidase, T7 polymerase, trp-lac (tac), trp E and the like.

An expression cassette for expression of synthase in a plant cell will include, in the 5' to 3' direction of transcription, a transcription and translation initiation control regulatory region (also known as a "promoter") functional in a plant cell, a nucleic acid sequence encoding a synthase, and a transcription termination region. Numerous transcription initiation regions are available which provide for a wide variety of constitutive or regulatable, e.g., inducible, transcription of the desaturase structural gene. Among transcriptional initiation regions used for plants are such regions associated with cauliflower mosaic viruses (35S, 19S), and structural genes such as for nopaline synthase or mannopine synthase or napin and ACP promoters, etc. The transcription/translation initiation regions corresponding to such structural genes are found immediately 5' upstream to the respective start codons. Thus, depending upon the intended use, different promoters may be desired.

Of special interest in this invention are the use of promoters which are capable of preferentially expressing the synthase in seed tissue, in particular, at early stages of seed oil formation. Selective modification of seed fatty acid/oils composition will reduce potential adverse effects to other plant tissues. Examples of such seed-specific promoters include the region immediately 5' upstream of a napin or seed ACP genes such as described in EP 0 255 378 (published 2/3/88), desaturase genes such as described in Thompson *et al* (*Proc. Nat. Acad. Sci.* (1991) *88:2578-2582)*, WO 92/03564 and Fig. 10 herein, or Bce-4 gene such as described in co-pending USSN 494,722, and Fig. 9 herein. Alternatively, the use of the 5' regulatory region associated with the plant synthase structural gene, i.e., the region immediately 5' upstream to a plant

synthase structural gene and/or the transcription termination regions found immediately 3' downstream to the plant synthase structural gene, may often be desired. In general, promoters will be selected based upon their expression profile which may change given the particular application.

Sequences found in an anti-sense orientation may be found in cassettes which at least provide for transcription of the sequence encoding the synthase. By anti-sense is meant a DNA sequence in the 5' to 3' direction of transcription which encodes a sequence complementary to the sequence of interest. It is preferred that an "anti-sense synthase" be complementary to a plant synthase gene indigenous to the plant host. Any promoter capable of expression in a plant host which causes initiation of high levels of transcription in all storage tissues during seed development is sufficient. Seed specific promoters may be desired.

A DNA sequence of this invention may include genomic or cDNA sequence. A cDNA sequence may or may not contain pre-processing sequences, such as transit peptide sequences. Transit peptide sequences facilitate the delivery of the protein to a given organelle and are cleaved from the amino acid moiety upon entry into the organelle, releasing the "mature" protein (or enzyme). As synthases are part of the FAS pathway of plastid organelles, such as the chloroplast, proplastid, etc., transit peptides may be required to direct the protein(s) to substrate. A transit peptide sequence from any plastid-translocating sources may be employed, such as from ACP, especially seed ACP, small subunit of ribulose bisphosphate carboxylase (RuBC), plant desaturase or from the native sequence naturally associated with the respective synthase.

The complete genomic sequence of a plant synthase may be obtained by the screening of a genomic library with a probe and isolating those sequences which hybridize thereto as described more fully below. Regulatory sequences immediately 5', transcriptional and translational initiation regions, and 3', transcriptional and

translational termination regions, to the synthase may be obtained and used with or without the synthase structural gene.

Other synthases and/or synthase nucleic sequences are obtainable from amino acid and DNA sequences provided herein. "Obtainable" refers to those plant synthases which have sufficiently similar sequence to that of the native sequence(s) of this invention to provide a biologically active synthase. One skilled in the art will readily recognize that antibody preparations, nucleic acid probes (DNA and RNA) and the like may be prepared and used to screen and recover synthases and/or synthase nucleic acid sequences from other sources. Thus, sequences which are homologously related to or derivations from either *R. communis* synthase I or II are considered obtainable from the present invention.

"Homologeously related" includes those nucleic acid sequences which are identical or conservatively substituted as compared to the native sequence. Typically, a homologously related nucleic acid sequence will show at least about 60% homology, and more preferably at least about 70% homology, between the *R. communis* synthase and the given plant synthase of interest, excluding any deletions which may be present. Homology is determined upon comparison of sequence information, nucleic acid or amino acid, or through hybridization reactions.

Probes can be considerably shorter than the entire sequence, but should be at least about 10, preferably at least about 15, more preferably at least 20 or so nucleotides in length. Longer oligonucleotides are also useful, up to the full length of the gene encoding the polypeptide of interest. Both DNA and RNA can be used.

A genomic library prepared from the plant source of interest may be probed with conserved sequences from a *R. communis* synthase cDNA to identify homologously related sequences. Use of an entire *R. communis* synthase cDNA may be employed if shorter probe sequences are not identified. Positive clones are then analyzed by restriction enzyme digestion and/or sequencing. In this general manner, one

or more sequences may be identified providing both the coding region, as well as the transcriptional regulatory elements of the synthase gene from such plant source. cDNA libraries prepared from other plant sources of interest may be screened as well, providing the coding region of synthase genes from such plant sources.

In use, probes are typically labeled in a detectable manner (for example with ^{32}P-labeled or biotinylated nucleotides) and are incubated with single-stranded DNA or RNA from the plant source in which the gene is sought, although unlabeled oligonucleotides are also useful. Hybridization is detected, typically using nitrocellulose paper or nylon membranes by means of the label after single-stranded and double-stranded (hybridized) DNA or DNA/RNA have been separated. Hybridization techniques suitable for use with oligonucleotides are well known to those skilled in the art. Thus, plant synthase genes may be isolated by various techniques from any convenient plant. Plant genes for synthases from developing seed obtained from other oilseed plants, such as *C. tinctorius* seed, rapeseed, cotton, corn, soybean cotyledons, jojoba nuts, coconut, peanuts, oil palm and the like are desired as well as from non-traditional oil sources, such as *S. oleracea* chloroplast, avocado mesocarp, *Cuphea*, California Bay and *Euglena gracillis*. Synthases, especially synthase I, obtained from *Cuphea* may show specialized activities towards medium chain fatty acids. Such synthase may be of special interest for use in conjunction with a plant medium-chain thioesterase.

Once the desired plant synthase sequence is obtained, it may be manipulated in a variety of ways. Where the sequence involves non-coding flanking regions, the flanking regions may be subjected to resection, mutagenesis, etc. Thus, transitions, transversions, deletions, and insertions may be performed on the naturally occurring sequence. In addition, all or part of the sequence may be synthesized, where one or more codons may be modified to provide for a modified amino acid sequence, or one or more codon mutations may be introduced to provide for a convenient

restriction site or other purpose involved with construction or expression. The structural gene may be further modified by employing synthetic adapters, linkers to introduce one or more convenient restriction sites, or the like. For expression, the open reading frame coding for the plant synthase or functional fragment thereof will be joined at its 5' end to a transcriptional initiation regulatory control region. In some instances, such as modulation of plant synthase via a nucleic acid sequence encoding synthase in an anti-sense orientation, a transcription initiation region or transcription/translation initiation region may be used. In embodiments wherein the expression of the synthase protein is desired in a plant host, a transcription/translation initiation regulatory region, is needed. Additionally, modified promoters, i.e., having transcription initiation regions derived from one gene source and translation initiation regions derived from a different gene source or enhanced promoters, such as double 35S CaMV promoters, may be employed for some applications.

As described above, of particular interest are those 5' upstream non-coding regions which are obtained from genes regulated during seed maturation, particularly those preferentially expressed in plant embryo tissue, such as ACP-and napin-derived transcription initiation control regions. Such regulatory regions are active during lipid accumulation and therefore offer potential for greater control and/or effectiveness to modify the production of plant desaturase and/or modification of the fatty acid composition. Especially of interest are transcription initiation regions which are preferentially expressed in seed tissue, i.e., which are undetectable in other plant parts. For this purpose, the transcript initiation region of acyl carrier protein isolated from *B. campestris* seed and designated as "Bcg 4-4" and a gene having an unknown function isolated from *B. campestris* seed and designated as "Bce-4" are also of substantial interest.

Briefly, Bce4 is found in immature embryo tissue at least as early as 11 days after anthesis (flowering),

peaking about 6 to 8 days later or 17-19 days post-anthesis, and becoming undetectable by 35 days post-anthesis. The timing of expression of the Bce4 gene closely follows that of lipid accumulation in seed tissue. Bce4 is primarily detected in seed embryo tissue and to a lesser extent found in the seed coat. Bce4 has not been detected in other plant tissues tested, root, stem and leaves.

Bce4 transcript initiation regions will contain at least 1 kb and more preferably about 5 to about 7.5 kb of sequence immediately 5' to the Bce4 structural gene.

The Bcg 4-4 ACP message presents a similar expression profile to that of Bce4 and, therefore, also corresponds to lipid accumulation in the seed tissue. Bcg 4-4 is not found in the seed coat and may show some differences in expression level, as compared to Bce4, when the Bcg 4-4 5' non-coding sequence is used to regulate transcription or transcription and translation of a plant Δ-9 desaturase of this invention.

The napin 1-2 message is found in early seed development and thus, also offers regulatory regions which can offer preferential transcriptional regulation of a desired DNA sequence of interest such as the plant desaturase DNA sequence of this invention during lipid accumulation. Napins are one of the two classes of storage proteins synthesized in developing *Brassica* embryos (Bhatty, *et al.*, *Can J. Biochem.* (1968) *46*:1191-1197) and have been used to direct tissue-specific expression when reintroduced into the *Brassica* genome (Radke, *et al.*, *Theor. Appl. Genet.* (1988) *75*:685-694).

As to regulatory transcript termination regions, these may be provided by the DNA sequence encoding the plant synthase or a convenient transcription termination region derived from a different gene source, especially the transcript termination region which is naturally associated with the transcript initiation region. Typically, the transcript termination region will contain at least about 1 kb, preferably about 3 kb of sequence 3' to the structural gene from which the termination region is derived.

In developing the DNA construct, the various components of the construct or fragments thereof will normally be inserted into a convenient cloning vector which is capable of replication in a bacterial host, e.g., *E. coli*. Numerous vectors exist that have been described in the literature. After each cloning, the plasmid may be isolated and subjected to further manipulation, such as restriction, insertion of new fragments, ligation, deletion, insertion, resection, etc., so as to tailor the components of the desired sequence. Once the construct has been completed, it may then be transferred to an appropriate vector for further manipulation in accordance with the manner of transformation of the host cell.

Normally, included with the DNA construct will be a structural gene having the necessary regulatory regions for expression in a host and providing for selection of transformed cells. The gene may provide for resistance to a cytotoxic agent, e.g. antibiotic, heavy metal, toxin, etc., complementation providing prototrophy to an auxotrophic host, viral immunity or the like. Depending upon the number of different host species into which the expression construct or components thereof are introduced, one or more markers may be employed, where different conditions for selection are used for the different hosts.

The manner in which the DNA construct is introduced into the plant host is not critical to this invention. Any method which provides for efficient transformation may be employed. Various methods for plant cell transformation include the use of Ti- or Ri-plasmids, microinjection, electroporation, liposome fusion, DNA bombardment or the like. In many instances, it will be desirable to have the construct bordered on one or both sides by T-DNA, particularly having the left and right borders, more particularly the right border. This is particularly useful when the construct uses *A. tumefaciens* or *A. rhizogenes* as a mode for transformation, although the T-DNA borders may find use with other modes of transformation.

Where *Agrobacterium* is used for plant cell transformation, a vector may be used which may be

introduced into the *Agrobacterium* host for homologous recombination with T-DNA or the Ti- or Ri-plasmid present in the *Agrobacterium* host. The Ti- or Ri-plasmid containing the T-DNA for recombination may be armed (capable of causing gall formation) or disarmed (incapable of causing gall formation), either being permissible, so long as the *vir* genes are present in the transformed *Agrobacterium* host. The armed plasmid can give a mixture of normal plant cell and gall.

A preferred method for the use of *Agrobacterium* as the vehicle for transformation of plant cells employs a vector having a broad host range replication system, at least one T-DNA boundary and the DNA sequence or sequences of interest. Commonly used vectors include pRK2 or derivatives thereof. See, for example, Ditta *et al.*, *PNAS USA*, (1980) 77:7347-7351 and EPA 0 120 515, which are incorporated herein by reference. Normally, the vector will be free of genes coding for opines, oncogenes and *vir*-genes. Included with the expression construct and the T-DNA will be one or more markers, which allow for selection of transformed *Agrobacterium* and transformed plant cells. A number of markers have been developed for use with plant cells, such as resistance to chloramphenicol, the aminoglycoside G418, hygromycin, or the like. The particular marker employed is not essential to this invention, one or another marker being preferred depending on the particular host and the manner of construction.

The vector is used for introducing the DNA of interest into a plant cell by transformation into an *Agrobacterium* having *vir*-genes functional for transferring T-DNA into a plant cell. The *Agrobacterium* containing the broad host range vector construct is then used to infect plant cells under appropriate conditions for transfer of the desired DNA into the plant host cell under conditions where replication and normal expression will occur. This will also usually include transfer of the marker, so that cells containing the desired DNA may be readily selected.

The expression constructs may be employed with a wide variety of plant life, particularly plant life involved in

the production of vegetable oils. These plants include, but are not limited to rapeseed, peanut, sunflower, *C. tinctorius*, cotton, *Cuphea*, soybean, and corn or palm.

For transformation of plant cells using *Agrobacterium*, explants may be combined and incubated with the transformed *Agrobacterium* for sufficient time for transformation, the bacteria killed, and the plant cells cultured in an appropriate selective medium. Once callus forms, shoot formation can be encouraged by employing the appropriate plant hormones in accordance with known methods and the shoots transferred to rooting medium for regeneration of plants. The plants may then be grown to seed and the seed used to establish repetitive generations and for isolation of vegetable oils.

The invention now being generally described, it will be more readily understood by reference to the following examples which are included for purposes of illustration only and are not intended to limit the present invention.

EXAMPLES

Example 1. Analysis of Synthase Proteins

In purification of synthase proteins as discussed in WO 92/03564, synthase II activity was observed only when both the 46 and 50kD peptides were present in *R. communis* protein preparations, while synthase I activity was detected in preparations containing only the 50kD peptide. In addition, *E. coli* expression data demonstrated that both the 46kD and 50kD synthase factors (factors A and B, respectively) were required for synthase II type activity, and that synthase factor A contributes the longer chain fatty acyl substrate specificity to synthase II activity.

To determine whether synthase II activity requires two discrete proteins or a single heterodimer, covalent intramolecular bonds are introduced into the purified protein preparation and the products of this reaction identified by SDS-PAGE and Western analysis. A similar analysis is conducted with a synthase I preparation to determine if synthase I activity is provided by a single peptide, or a homodimer of the observed 50kD peptide.

Twenty µg (400pmol of 46kD and 50kD peptides) of a purified synthase II preparation in "20 buffer", is combined with 40nmol of EGS (ethylene glycol *bis*(succinimidyl succinate)) in 10% Me_2SO, to a final volume of 0.4ml. The reaction is stopped by addition of 0.045ml of 2M Tris-HCl, pH 7.0, and the protein is prepared immediately for gel electrophoresis by the addition of SDS-PAGE sample buffer containing ß-mercaptoethanol. A purified preparation of synthase I is similarly treated, except that 600ng (12pmol of 50kD peptide) is combined with 8.3nmol of EGS in a volume of 85 µl. This reaction is stopped by addition of 9µl of 2M Tris-HCl, pH 7.0. The crosslinked proteins are analyzed by SDS-PAGE, Western transfer and antibody blotting.

Each of the two active sites on a molecule of EGS can form a covalent bond with any available amino group of the studied proteins, resulting in linkage of the two amino groups across an EGS bridge. With synthase II, at room temperature incubations of 10 or 30 minutes in from 0.1mM to 10mM EGS, only two major and two minor species of crosslinked proteins are formed. These proteins are observed to migrate on SDS-PAGE at about 124 and 107kD, suggesting that the proteins are dimers. By Western analysis, all crosslinked products react positively with antibodies raised against both the 46 and 50kD peptides, indicating that both peptides are present in all products of crosslinking. The appearance of more than one dimer could reflect different conformations of the dimers depending on the number and the locations of intramolecular bond formations. Only after prolonged incubation, or with higher concentrations of EGS, are multimeric protein species formed. These results provide additional evidence that the synthase II protein is a heterodimer of the 46 and 50kD subunits.

When synthase I protein is subjected to the same reactions, one major and two minor products are formed, each of which has a mobility of about 116kD. All three of these products react with antibody raised against the 50kD

peptide, but not with antibodies raised against the 46kD peptide. These results suggest that the protein having synthase I activity is a homodimer of the 50kD peptide.

Example 2. Synthase Gene Sequence

The preparation of a cDNA libraries, using the methods as described in Alexander, *et al.* (*Methods in Enzymology* (1987) *154*:41-64), and the screening of the cDNA libraries for synthase cDNA clones are described in WO 92/03564. Sequences of plant synthase factor proteins A and B which have been shown to be required for synthase II activitiy are provided herein in Figures 1-8. It is noted that synthase factor B protein is also required for plant synthase I type activity (WO 92/03564)

Sequence of an *E. coli* synthase III gene is found in Tsay *et. al.* (*J. Biol. Chem.* (1992) *267*:6807-6814)

Example 3. Expression Cassettes

In this example, expression cassettes suitable for insertion of synthase genes are described.

Expression cassettes utilizing 5'-upstream sequences and 3'-downstream sequences of genes preferentially expressed during seed development can be constructed from isolated DNA sequences of genes with an appropriate expression pattern. Examples of genes which are expressed during seed development in *Brassica* are a napin gene, 1-2, and an ACP gene, Bcg4-4, both described in European Patent Publication EP 0 255 378, and a Bce4 gene, as described below. The napin gene encodes a seed storage protein that is preferentially expressed in immature embryos which are actively producing storage proteins. The ACP gene encodes a protein which is an integral factor in the synthesis of fatty acids in the developing embryo and is preferentially expressed during fatty acid synthesis. Bce4 is a gene that produces a protein of unknown function that is preferentially expressed early in embryo development, at about 15-19 days post-anthesis, and is also detectable as early as 11 days post-anthesis. The sequence of Bce4 is shown in Figure 9.

DNA sequences that control the expression of these genes can be isolated and sufficient portions of the 5' and 3' regulatory regions combined such that a synthase gene inserted between these sequences will be preferentially expressed early in seed development. This expression pattern will allow the synthase gene to affect fatty acid synthesis, which also occurs early in seed development. For example, a 1.45 kb *Xho*I fragment containing 5' sequence and a 1.5 kb *Sst*I/*Bgl*II fragment containing 3' sequence of the Bcg4-4 ACP gene can be combined in an ACP expression cassette using a variety of available DNA manipulation techniques. Similarly, a napin expression cassette can be prepared that contains approximately 1.725 kb of 5' sequence from an EcoRV site to immediately before the ATG start codon and approximately 1.25 kb of 3' sequence from an XhoI site approximately 18 bases past the TAG stop codon to a 3' HindIII site of a 1-2 napin gene. A Bce4 expression cassette can be made by combining approximately 7.4 kb of 5' DNA sequence from an upstream *Pst*I site to immediately before the ATG start codon with approximately 1.9 kb of 3' sequences from immediately after the TAA stop codon to a 3' *Pst*I site.

Variations can be made in these expression cassettes such as increasing or decreasing the amounts of 5' and 3' sequences, combining the 5' sequences of one gene with the 3' sequences of a different gene (for example using the 1.3 kb 5' sequences of napin 1-2 with the 1.5 kb 3' sequences of ACP Bcg4-4 in an expression cassette), or using other available 3' regulatory sequences, as long as these variations result in expression cassettes that allow for expression of the inserted synthase gene at an appropriate time during seed development.

A. <u>Napin Seed Specific Expression Cassettes</u>

1. Napin 1-2 pCGN1808 Expression Cassette

An expression cassette utilizing 5' upstream sequences and 3' downstream sequences obtainable from *B. campestris* napin gene can be constructed as follows.

A 2.7 kb *Xho*I fragment of napin 1-2 (*See*, Figure 2 of EP 0 255 378, published February 3, 1988) containing 5'

upstream sequences is subcloned into pCGN789 (a pUC based vector with the normal polylinker replaced by the synthetic linker which encodes the restriction digestion sites *Eco*RI, *Sal*I, *Bgl*II, *Pst*I, *Xho*I, *Bam*HI, *Hin*dIII) and results in pCGN940. The majority of the napin coding region of pCGN940 was deleted by digestion with *Sal*I and religation to form pCGN1800. Single-stranded DNA from pCGN1800 was used in an *in vitro* mutagenesis reaction (Adelman *et al.*, *DNA* (1983) *2*:183-193) using a synthetic oligonucleotide which inserted *Eco*RV and *Nco*I restriction site at the junction of the promoter region and the ATG start codon of the napin gene. An appropriate mutant was identified by hybridization to the oligonucleotide used for the mutagenesis and sequence analysis and named pCGN1801.

A 1.7 kb promoter fragment was subcloned from pCGN1801 by partial digestion with *Eco*RV and ligation to pCGN786 (a pCGN566 (polylinker in opposite orientation as pCGN565 described in WO 92/03564) chloramphenicol based vector with the synthetic linker described above in place of the normal polylinker) cut with *Eco*RI and blunted by filling in with DNA Polymerase I Klenow fragment to create pCGN1802.

A 2.1 kb *Sal*I fragment of napin 1-2 containing 3' downstream sequences is subcloned into pCGN789 (described above) and results in pCGN941. pCGN941 is digested with *Xho*I and *Hin*dIII and the resulting approximately 1.6 kb of napin 3' sequences are inserted into *Xho*I-*Hin*dIII digested pCGN1802 to result in pCGN1803. In order to remove a 326 nucleotide *Hin*dIII fragment inserted opposite to its natural orientation, as a result of the fact that there are 2 *Hin*dIII sites in pCGN1803, the pCGN1803 is digested with *Hin*dIII and religated. Following religation, a clone is selected which now contains only 1.25 kb of the original 1.6 napin 3' sequence. This clone, pCGN1808 is the napin 1-2 expression cassette and contains 1.725 kb of napin promoter sequences and 1.265 kb of napin 3' sequence with the unique cloning sites *Sal*I, *Bgl*I, *Pst*I and *Xho*I in between.

2. Napin 1-2 pCGN3223 Expression Cassette

Alternatively, pCGN1808 may be modified to contain flanking restriction sites to allow movement of only the expression sequences and not the antibiotic resistance marker to binary vectors such as pCGN1557 (McBride and Summerfelt, *supra*). Synthetic oligonucleotides containing *Kpn*I, *Not*I and *Hin*dIII restriction sites are annealed and ligated at the unique *Hin*dIII site of pCGN1808, such that only one *Hin*dIII site is recovered. The resulting plasmid, pCGN3200 contains unique *Hin*dIII, *Not*I and *Kpn*I restriction sites at the 3'-end of the napin 3'-regulatory sequences as confirmed by sequence analysis.

The majority of the napin expression cassette is subcloned from pCGN3200 by digestion with *Hin*dIII and *Sac*I and ligation to *Hin*dIII and *Sac*I digested pIC19R (Marsh, *et al.* (1984) *Gene 32*:481-485) to make pCGN3212. The extreme 5'-sequences of the napin promoter region are reconstructed by PCR using pCGN3200 as a template and two primers flanking the *Sac*I site and the junction of the napin 5'-promoter and the pUC backbone of pCGN3200 from the pCGN1808 construct. The forward primer contains *Cla*I, *Hin*dIII, *Not*I, and *Kpn*I restriction sites as well as nucleotides 408-423 of the napin 5'-sequence (from the *Eco*RV site) and the reverse primer contains the complement to napin sequences 718-739 which include the unique *Sac*I site in the 5'-promoter. The PCR was performed using a Perkin Elmer/Cetus thermocycler according to manufacturer's specifications. The PCR fragment is subcloned as a blunt-ended fragment into pUC8 (Vieira and Messing (1982) *Gene 19*:259-268) and digested with *Hinc*II to give pCGN3217. Sequence of pCGN3217 across the napin insert verifies that no improper nucleotides were introduced by PCR. The napin 5-sequences in pCGN3217 are ligated to the remainder of the napin expression cassette by digestion with *Cla*I and *Sac*I and ligation to pCGN3212 digested with *Cla*I and *Sac*I. The resulting expression cassette pCGN3221, is digested with *Hin*dIII and the napin expression sequences are gel purified away and ligated to pIC20H (Marsh, *supra*) digested with *Hin*dIII. The final expression cassette is pCGN3223, which contains in an ampicillin resistant background, essentially

identical 1.725 napin 5' and 1.265 3' regulatory sequences as found in pCGN1808. The regulatory regions are flanked with *Hin*dIII, *Not*I and *Kpn*I restriction sites and unique *Sal*I, *Bgl*II, *Pst*I, and *Xho*I cloning sites are located between the 5' and 3' noncoding regions.

B. <u>Bce4 Expression Cassette</u>

An expression cassette for seed specific expression can also be constructed from Bce4 gene sequences, such as those represented in Figure 9. Genomic clones having regulatory sequences of the Bce4 gene may be isolated from a *Brassica campestris* genomic library using Bce4 sequences as probe. For example, an approximately 20 kb *Bam*HI fragment is isolated and designated as clone P1C1. The approximately 20 kb insert of clone P1C1 is released by *Bam*HI digestion and inserted into the *Bam*HI site of the binary vector pCGN1547 (see below), producing pCGN1853. The *Pst*I fragment of pCGN1853, containing the Bce4 gene, is inserted into the *Pst*I site of pUC18 (Norrander, *et al.* (1983) *Gene 26*:101-106), producing pCGN1857. The plasmid pCGN1857 was deposited with the ATCC, Rockville, MD on March 9, 1990, accession number 68251. The *Cla*I fragment of pCGN1857, containing the Bce4 gene is ligated into *Cla*I digested Bluescript KS+ (Stratagene; La Jolla, CA), producing pCGN1864. Single stranded DNA is made from pCGN1864 and altered by *in vitro* mutagenesis as described by Adelman *et al.* (*DNA* (1983) *2*:183-193) using oligonucleotides having homology to Bce4 sequences 5' and 3' of the translated start and stop codons and also coding for restriction digest sites. The resulting plasmid, pCGN1866, contains *Xho*I and *Bam*HI sites (from BCE45P) immediately 5' to the Bce4 start codon and *Bam*HI and *Sma*I sites (from BCE43P) immediately 3' to the Bce4 stop codon. The *Cla*I fragment of pCGN1866, containing the mutagenized sequences, is inserted into the *Cla*I site of pCGN2016 (described below), producing pCGN1866C. The *Cla*I fragment of pCGN1866C is used to replace the corresponding wild-type *Cla*I fragment of pCGN1867 (described below) to produce pCGN1868. Bce4 coding sequences are removed by digestion of pCGN1868 with *Bam*HI and recircularization of the plasmid

to produce pCGN1870. The Bce4 expression cassette, pCGN1870, contains 7.4 kb of 5' regulatory sequence and 1.9 kb of 3' regulatory sequence derived from the Bce4 genomic clone separated by the cloning sites, *Xho*I, *Bam*HI, and *Sma*I.

pCGN1867

The *Bam*HI and *Sma*I sites of pUC18 (Norrander *et al.*, (1983) *supra*) are removed by *Bam*HI-*Sma*I digestion and recircularization of the plasmid, without repair of the ends, to produce pCGN1862. The *Pst*I fragment of pCGN1857, containing the Bce4 gene, is inserted into the *Pst*I site of pCGN1862 to produce pCGN1867.

pCGN2016

The multiple cloning sites of pUC12-Cm (Buckley, K., Ph.D. Thesis, UCSD, CA (1985)) are replaced by those of pUC18 to produce pCGN565. The *Hha*I fragment of pCGN565, containing the chloramphenicol resistance gene is excised, blunted by use of mung bean nuclease, and inserted into the *Eco*RV site of Bluescript KS- (Stratagene; La Jolla, CA) to create pCGN2008. The chloramphenicol resistance gene of pCGN2008 is removed by *Eco*RI-*Hin*dIII digestion. After treatment with Klenow enzyme to blunt the ends, the fragment carrying the chloramphenicol resistance gene is inserted into the *Dra*I site of Bluescript KS-, replacing the ampicillin resistance gene of Bluescript KS-, to produce pCGN2016.

C. ACP Expression Cassette

An expression cassette utilizing 5'-upstream sequences and 3'-downstream sequences obtainable from *B. campestris* ACP gene can be constructed as follows. A 1.45kb *Xho*I fragment of Bcg 4-4 containing 5'-upstream sequences is subcloned into the cloning/sequencing vector Bluescript+ (Stratagene Cloning Systems, San Diego, CA). The resulting construct, pCGN1941, is digested with *Xho*I and ligated to a chloramphenicol resistant Bluescript M13+ vector, pCGN2015 digested with *Xho*I. pCGN2015 is described in WO 92/05364. This alters the antibiotic resistance of the plasmid from penicillin resistance to chloramphenicol resistance. The chloramphenicol resistant plasmid is pCGN1953.

3'-sequences of Bcg 4-4 are contained on an *Sst*I/BglII fragment cloned in the *Sst*I/*Bam*HI sites of M13 Bluescript+ vector. This plasmid is named pCGN1940. pCGN1940 is modified by *in vitro* site-directed mutagenesis (Adelman *et al.*, *DNA* (1983) *2*:183-193) using the synthetic oligonucleotide (SEQ ID NO:51) 5'-CTTAAGAAGTAACCCGGGCTGCAGTTTTAGTATTAAGAG-3' to insert *Sma*I and *Pst*I restriction sites immediately following the stop codon of the reading frame for the ACP gene 18 nucleotides from the *Sst*I site. The 3'-noncoding sequences from this modified plasmid, pCGN1950, are moved as a *PsI-Sma*I fragment into pCGN1953 cut with *Pst*I and *Sma*I. The resulting plasmid pCGN1977 comprises the ACP expression cassette with the unique restriction sites *Eco*RV, *Eco*RI and *Pst*I available between the 1.45kb 5' and 1.5 kb of 3'-noncoding sequences for the cloning of genes to be expressed under regulation of these ACP gene regions.

Example 4. Synthase Constructs

A. Preparation of Plant Transformation Vectors

Synthase cDNA sequences can be inserted in expression cassettes containing plant regulatory regions using a variety of DNA manipulation techniques. In this manner, synthase constructs in either the sense or anti-sense orientation are prepared. If convenient restriction sites are present in the synthase clones, they may be inserted into the expression cassette by digesting with the restriction endonucleases and ligation into the cassette that has been digested at one or more of the available cloning sites. If convenient restriction sites are not available in the clones, the DNA of either the cassette or the synthase gene(s), can be modified in a variety of ways to facilitate cloning of the synthase gene(s) into the cassette. Examples of methods to modify the DNA include by PCR, synthetic linker or adaptor ligation, *in vitro* site-directed mutagenesis (Adelman *et al.*, supra), filling in or cutting back of overhanging 5' or 3' ends, and the like. These and other methods of manipulating DNA are well known to those of ordinary skill in the art.

The fragment containing the synthase gene in the expression cassette, 5' sequences/synthase/3' sequences, is then cloned into a binary vector, such as described by McBride and Summerfelt (*Pl.Mol.Biol.* (1990) *14*:269-276), for *Agrobacterium* transformation. Other binary vectors are known in the art and may also be used for synthase cassettes. The binary vector containing the expression cassette and the synthase gene is transformed into *Agrobacterium tumefaciens*, such as strain EHA101 (Hood, *et al.*, *J. Bacteriol.* (1986) *168*:1291-1301) as per the method of Holsters, *et al.*, (*Mol. Gen. Genet.* (1978) *163*:181-187), and used to generate transformed plants as described in Example 5.

B. <u>Synthase Factor A Constructs</u>

1. Sense orientation

Constructs containing sense synthase sequences under the control of plant regulatory regions for expression in plant cells may be prepared as follows. The *R. communis* synthase factor A cDNA, 1-1A, is altered by *in vitro* mutagenesis to insert a *Bam*HI restriction site at the 5' end of the cDNA insert and *Xho*I and *Sma*I sites immediately 3' of the translation stop codon. The resulting construct, pCGN2781, is digested with *Bam*HI and *Xho*I and ligated into *Bgl*II and *Xho*I digested pCGN3223, the above described napin expression cassette, resulting in pCGN2785. The napin/factor A/napin region of pCGN2785 is obtained by digestion with *Asp*718 and ligated into *Asp*718 digested pCGN1557 (McBride *et al.*; *supra*), resulting in pCGN2787.

2. Antisense Orientation.

For antisense synthase A, a *Bam*HI/*Eco*RV fragment of the *B. campestris* synthase factor A cDNA clone, pCGN4300 (nucleotides 218-1535), is treated to create blunt ends, and subcloned into the *Eco*RV site of the ACP expression cassette, pCGN1977, to create pCGN4304. A *Kpn*I/*Xba*I fragment containing the ACP 5'/antisense *Brassica* factor A/ACP 3' fragment is inserted into *Kpn*I/*Xba*I digested binary vector pCGN1557 (McBride and Summerfelt, *supra*) resulting in plant transformation construct pCGN4306.

To insert the *Brassica* synthase factor A cDNA clone into a napin cassette, the *Bam*HI/*Eco*RV fragment of pCGN4300 (nucleotides 218-1535) is blunted and ligated into the *Bgl*II-digested napin cassette, pCGN3223, which has been similarly treated to provide a blunt ended DNA molecule. The resulting plasmid is pCGN4305. An *Asp*718 fragment containing the napin 5'/antisense *Brassica* A/napin 3' fragment is subcloned into binary vector pCGN1557 to produce a construct for plant transformation, pCGN4324.

3. Transit Fusion Constructs.

Due to the possibility that the *R. communis* factor A cDNA clone does not encode the entire transit peptide, DNA constructs are prepared to fuse the transit peptide encoding region from the *Brassica* factor A cDNA clone (including the V-A-A-C-M-S conserved region) to the mature peptide encoding region from the *R. communis* factor A clone. The constructs are designed such that the encoding region for the transit peptide and first 24 amino acids of the mature *R. communis* protein (lysine residue at nucleotides 365-367 of sequence shown in Figure 3 is the presumed N-terminus) is replaced with the corresponding region from the *Brassica* clone.

Two different 5' regions of the *Brassica* factor A cDNA are obtained by PCR. The "short" version includes nucleotides 79-423 of the sequence provided in Figure 6. This region encodes from the first methionine residue of the *Brassica* factor A cDNA to the histidine residue at position 115. The "long" version includes nucleotides 7-423 of the sequence shown in Figure 6, and thus includes a portion of the factor A 5' untranslated region. Each of these *Brassica* synthase factor A fragments also contains *Sal*I and *Eco*RV restriction sites at their 5' and 3' ends, respectively, which were provided in the oligonucleotide primers used in the PCR. The insertion of the 3' *Eco*RV site alters the codon for the 116 Asp from "GAC" to "GAT".

An *R. communis* synthase factor A DNA fragment, containing the encoding region for amino acids 145-540 of the sequence shown in Figure 3, and a stop codon, is also obtained by PCR. The coding sequence "GATATC" is chosen

for the *Asp* and *Ile* amino acids at positions 145-146 in order to provide an *Eco*RV site at the 5' end of this fragment. The PCR primers are designed such that an *Xho*I site is inserted at the 3' end of this fragment immediately following the "TGA" stop codon.

The *Brassica* and *R. communis* synthase factor A fragments are fused by ligation of their respective 3' and 5' *Eco*RV sites. The fusion fragment is obtained by digestion with SalI and XhoI and ligated into the napin expression cassette, pCGN3223. The construct containing the "short" *Brassica* factor A is designated pCGN4313, and the construct containing the "long" *Brassica* factor A is designated pCGN4314. The *Kpn*I fragments containing the napin 5'/synthase A fusion/napin 3' fragments are subcloned into binary vector pCGN1557 to produce vectors for plant transformation. The "short" *Brassica* factor A construct is designated pCGN4319, and the "long" *Brassica* factor A construct is designated pCGN4320.

A third fusion construct with the mature *R. communis* synthase factor A is prepared which incorporates the transit peptide encoding sequence of the *C. tinctorius* desaturase shown in Figure 10. A DNA fragment encoding amino acids 121-540 of Figure 3 (the mature *R. communis* synthase factor A and the asparagine residue immediately N-terminal to the mature peptide) is obtained by PCR. In addition, the sequence, CCATGGCC is included in the forward oligonucleotide primer and added at the 5' terminus of this fragment, such that an *Nco*I restriction site precedes the synthase sequence, and methionine and alanine residues are added to the synthase peptide encoding region. An *Xho*I restriction site is provided immediately following the stop codon at the 3' terminus.

The *C. tinctorius* desaturase cDNA clone shown in Figure 10, pCGN2754, is modified by PCR to insert *Pst*I, *Sma*I and *Xho*I sites to flank the coding region. The PCR product is digested with *Pst*I and ligated to pUC8 (Vieira and Messing (1982) *Gene 19*:2359-268) digested with *Pst*I to produce pCGN3220. The large *Nco*I/*Sac*I fragment of pCGN3220 containing the pUC8 vector and the *C. tinctorius* desaturase

cDNA sequence 5' to the *Nco*I site and 3' to the *Sac*I site is gel purified and ligated to the gel-purified *Nco*I/*Sac*I internal fragment of pCGN2754 resulting in pCGN3222. The coding region of the *C. tinctorius* desaturase from pCGN3222 is cloned into the pCGN3223 napin cassette by digestion with *Xho*I and ligation to pCGN3223 digested with *Xho*I and *Sal*I, resulting in pCGN3229.

pCGN3229 is digested with *Nco*I and *Xho*I to remove the mature desaturase encoding region. The *R. communis* synthase factor A fragment described above is digested with *Nco*I and *Xho*I and ligated to the *Nco*I/*Xho*I digested pCGN3229. This results in pCGN4308, the napin 5'/desaturase:synthase A/napin 3' fusion construct. pCGN4308 is digested with *Asp*718 and subcloned into binary vector pCGN1557 to produce plant transformation construct pCGN4318.

C. <u>Synthase III Constructs</u>

Fusion constructs of the bacterial synthase III encoding sequence and various plant transit peptide encoding sequences may be prepared. These constructs are used for generation of transgenic plants, wherein the bacterial synthase is incorporated into the chloroplasts for interaction with the plant fatty acid synthesis enzymes.

A fusion of the *Brassica* ACP transit peptide encoding sequence from a *B. rapa* (formerly *campestris*) seed ACP cDNA (Rose *et al.* (1987) *Nuc. Acids Res. 15*:7197) and the ß-ketoacyl-acyl carrier protein synthase III gene (*fab*H) from *E. coli* K-12 (Tsay *et al.* (1992) *J. Biol Chem. 267*:6807-6814), is prepared as follows. The *B. rapa* ACP transit peptide encoding region plus the 5' untranslated sequence is obtained by PCR, wherein the oligonucleotide primers are designed such that an *Bam*HI site is added immediately 5' to the *Xho*I site at the 5' end of the *B. rapa* cDNA clone, and an *Nhe*I site is inserted immediately 3' to the cysteine codon at the 3' end of the transit peptide encoding region. The *fab*H encoding region is obtained by PCR from *E. coli* DNA, with oligonucleotide primers designed such that an *Nhe*I site is inserted immediately 5' to the N-terminal

methionine codon, and *Xho*I and *Sma*I sites are inserted immediately 3' to the TAG stop codon. The *Nhe*I site adds an alanine and serine encoding region immediately 5' to the *fab*H N-terminal methionine.

The ACP and synthase III fragments are obtained by ligation at the inserted *Nhe*I restriction sites. The ACP/synthase III fusion fragment is inserted into an appropriate cassette containing plant regulatory regions. For example, for regulation under the napin regulatory regions, the ACP/synthase III fragment is obtained by digestion with *Bgl*II and *Xho*I and ligated into pCGN3223. Similarly, constructs wherein the fusion ACP/synthase III fragment is positioned for control under various other plant regulatory regions may be obtained. Other regulatory regions of interest include the Bce4 and ACP regions for seed expression, as well as 35S, double 35S, or T-DNA promoter regions (such as *mas* and *nos*), to provide for constitutive expression in various plant tissues. Constitutive expression may be desirable to test for uptake into chloroplasts of the synthase protein produced by such constructs, for example by electroporation into plant protoplasts and Western analysis. Constitutive expression is also useful for analysis of effects of the expression of synthase in various plant tissues, such as leaves, roots and stems.

Additional ACP/synthase III fusion constructs may be prepared which include various portions of the ACP mature protein encoding region in addition to the ACP transit peptide encoding region. For example, a fusion containing the *B. rapa* ACP transit encoding sequence plus coding sequence for an additional 12 amino acids of the mature ACP protein is prepared. The *fab*H encoding region is obtained by PCR from *E. coli* DNA, with oligonucleotide primers designed such that an *DdeI* site, "CTAAG" is inserted immediately 5' to the N-terminal methionine codon, and *Xho*I and *Sma*I sites are inserted immediately 3' to the TAG stop codon. The *B. rapa* ACP clone contains a *DdeI* site within the codons for amino acids 11-12 (Ser-Lys) of the mature protein region. Thus, the *B. rapa* ACP transit plus 12

fragment is obtained by digestion with *Dde*I and an appropriate site 5' to the ATG start codon. This fragment is ligated to the synthase III fragment at the *Dde*I site to form the ACP transit + 12/synthase III fusion. This fusion is inserted into a napin expression cassette by digestion with *Xho*I and ligation to *Xho*I digested pCGN3223. As discussed above, additional constructs for transcriptional control of the synthase III fusion under various other plant regulatory elements may be similarly prepared.

In addition to the ACP transit peptide discussed above, various other plant transit peptides are known in the art, and may be used in a similar manner. For example, the *Brassica* synthase A transit peptide used above in *R. communis* synthase fusion constructs, may also find use in conjunction with the bacterial synthase III. Similarly, other known transit peptides, such as those for SSU, stearoyl ACP desaturase and other nuclear encoded chloroplast proteins may be substituted for the ACP transit peptide.

D. <u>Constructs Containing More than One Synthase Gene</u>

If more than one synthase gene is required to obtain an optimum effect in plants, the genes may be expressed under regulation of the same promoter, or alternatively under regulation of two different promoters that are preferentially expressed in developing seeds, such as the napin, ACP, and Bce4 sequences described above. The constructs may then be introduced into plants in the same binary vector, or introduced simultaneously in different binary vectors.

For example, for expression of both synthase factors A and B in plant cells, a construct is prepared where *R. communis* synthase factor A and *R. communis* synthase factor B genes are each under the control of napin regulatory regions in the same binary vector.

The napin/factor A/napin region of pCGN2785 is obtained by digestion with *Asp*718 and ligated into *Asp*718 digested pCGN1557 (McBride *et al.*; *supra*), resulting in pCGN2787. pCGN2787 is digested at the unique *Pst*I site and treated with T4 polymerase to fill in the 3' overhang, and

digested with calf intestinal alkaline phosphatase to dephosphorylate the 5' termini to prevent self-ligation of pCGN2787. The napin/factor B/napin region of pCGN2786 is obtained by digestion with *Hin*dIII and the Klenow fragment of DNA polymerase to provide a blunt-ended DNA fragment, which is then ligated to the T4 polymerase blunt-ended pCGN2787 DNA. The resulting construct, pCGN2797, contains the *R. communis* synthase factors A and B, each positioned for expression from a napin promoter region.

Example 5. Plant Transformation

In this example, an *Agrobacterium*-mediated plant transformation is described and *Brassica napus* is exemplified. Also, a DNA-bombardment plant transformation is described and peanut transformation is exemplified.

A. <u>*Agrobacterium* Mediated Transformation</u>

Transformation of *Brassica* species is described by Radke *et al.* (*Theor. Appl. Genet.*(1988) *75*:685-694; *Plant Cell Reports* (1992) *11*:499-505). Seeds of *Brassica napus* cv. Delta are soaked in 95% ethanol for 2 min, surface sterilized in a 1.0% solution of sodium hypochlorite containing a drop of Tween 20 for 45 min., and rinsed three times in sterile, distilled water. Seeds are then plated in Magenta boxes with 1/10th concentration of Murashige minimal organics medium (Gibco) supplemented with pyrodoxine (50 μg/l), nicotinic acid (50 μg/l), glycine (200 μg/l), and 0.6% Phytagar (Gibco) pH 5.8. Seeds are germinated in a culture room at 22°C in a 16 h photoperiod with cool fluorescent and red light of intensity approximately 65 μEinsteins per square meter per second ($\mu Em^{-2}s^{-1}$).

Hypocotyls are excised from 7 day old seedlings, cut into pieces approximately 4 mm in length, and plated on feeder plates (Horsch et al. 1985). Feeder plates are prepared one day before use by plating 1.0 ml of a tobacco suspension culture onto a petri plate (100x25 mm) containing about 30 ml MS salt base (Carolina Biological) 100 mg/l inositol, 1.3 mg/l thiamine-HCl, 200 mg KH_2PO_4 with 3% sucrose, 2,4-D (1.0 mg/l), 0.6% Phytagar, and pH

adjusted to 5.8 prior to autoclaving (MS0/1/0 medium). A sterile filter paper disc (Whatman 3 mm) is placed on top of the feeder layer prior to use. Tobacco suspension cultures are subcultured weekly by transfer of 10 ml of culture into 100 ml fresh MS medium as described for the feeder plates with 2,4-D (0.2 mg/l), Kinetin (0.1 mg/l). All hypocotyl explants are preincubated on feeder plates for 24 h. at 22°C in continuous light of intensity 30 $\mu Em^{-2}S^{-1}$ to 65 $\mu EM^{-2}S^{-1}$.

Single colonies of *A. tumefaciens* strain EHA101 containing a binary plasmid are transferred to 5 ml MG/L broth and grown overnight at 30°C. Per liter, MG/L broth contains 5g mannitol, 1 g L-glutamic acid or 1.15 g sodium glutamate, 0.25 g kH_2PO_4, 0.10 g NaCL, 0.10 g $MGSO_4 \cdot 7H_2O$, 1 mg biotin, 5 g tryptone, and 2.5 g yeast extract, and the broth is adjusted to pH 7.0. Hypocotyl explants are immersed in 7-12 ml MG/L broth with bacteria diluted to $1x10^8$ bacteria/ml and after 10-20 min. are placed onto feeder plates. After 48 h of co-incubation with *Agrobacterium*, the hypocotyl explants are transferred to B5 0/1/0 callus induction medium which contains filter sterilized carbenicillin (500 mg/l, added after autoclaving) and kanamycin sulfate (Boehringer Mannheim) at concentrations of 25 mg/l.

After 3-7 days in culture at 65 $\mu Em^{-2}S^{-1}$ to 75 $\mu Em^{-2}S^{-1}$ continuous light, callus tissue is visible on the cut surface and the hypocotyl explants are transferred to shoot induction medium, B5BZ (B5 salts and vitamins supplemented with 3 mg/l benzylaminopurine, 1 mg/l zeatin, 1% sucrose, 0.6% Phytagar and pH adjusted to 5.8). This medium also contains carbenicillin (500 mg/l) and kanamycin sulfate (25 mg/l). Hypocotyl explants are subcultured onto fresh shoot induction medium every two weeks.

Shoots regenerate from the hypocotyl calli after one to three months. Green shoots at least 1 cm tall are excised from the calli and placed on medium containing B5 salts and vitamins, 1% sucrose, carbenicillin (300 mg/l), kanamycin sulfate (50 mg/l) and 0.6% Phytagar) and placed in a culture room with conditions as described for seed

germination. After 2-4 weeks shoots which remain green are cut at the base and transferred to Magenta boxes containing root induction medium (B5 salts and vitamins, 1% sucrose, 2 mg/l indolebutyric acid, 50 mg/l kanamycin sulfate and 0.6% Phytagar). Green rooted shoots are tested for NPT II activity.

Transgenic *Arabidopsis thaliana* plants may also be obtained by *Agrobacterium*-mediated transformation using similar techniques. For example, a useful method has been described by Valverkens *et al.*, (*Proc. Nat. Acad. Sci.* (1988) *85*:5536-5540).

B. <u>Transformation by Particle Bombardment</u>

DNA sequences of interest may be introduced as expression cassettes, comprising at least a promoter region, a gene of interest, and a termination region, into a plant genome via particle bombardment as described in European Patent Application 332 855 and in co-pending application USSN 07/225,332, filed July 27, 1988.

Briefly, tungsten or gold particles of a size ranging from 0.5µM-3µM are coated with DNA of an expression cassette. This DNA may be in the form of an aqueous mixture or a dry DNA/particle precipitate.

Tissue used as the target for bombardment may be from cotyledonary explants, shoot meristems, immature leaflets, or anthers.

The bombardment of the tissue with the DNA-coated particles is carried out using a Biolistics™ particle gun (Dupont; Wilmington, DE). The particles are placed in the barrel at variable distances ranging from 1cm-14cm from the barrel mouth. The tissue to be bombarded is placed beneath the stopping plate; testing is performed on the tissue at distances up to 20 cm. At the moment of discharge, the tissue is protected by a nylon net or a combination of nylon nets with mesh ranging from 10µM to 300µM.

Following bombardment, plants may be regenerated following the method of Atreya, *et al.*, (*Plant Science Letters* (1984) *34*:379-383). Briefly, embryo axis tissue or cotyledon segments are placed on MS medium (Murashige and Skoog, *Physio. Plant.* (1962) *15*:473) (MS plus 2.0 mg.l 6-

benzyladenine (BA) for the cotyledon segments) and incubated in the dark for 1 week at 25 ± 2°C and are subsequently transferred to continuous cool white fluorescent light (6.8 W/m^2). On the 10th day of culture, the plantlets are transferred to pots containing sterile soil, are kept in the shade for 3-5 days are and finally moved to greenhouse.

The putative transgenic shoots are rooted. Integration of exogenous DNA into the plant genome may be confirmed by various methods known to those skilled in the art.

Example 6. Analysis of Transformed Plants

Seeds from 15 *Arabidopsis* plants transformed with pCGN2797 (napin 5'/*R. communis* synthase factor A/napin 3' and napin 5'/*R. communis* synthase factor B/napin 3') were analyzed for the presence of *R. communis* synthase proteins. Five of these plants test positive, by Western analysis, for expression of the 50kD *R. communis* synthase factor B protein. Cross-reactivity of the *R. communis* synthase factor A polyclonal antibody with the corresponding *Brassica* synthase protein, prevents detection of this synthase protein by Western analysis.

Two of the plants which tested positive for expression of the 50kD *R. communis* synthase protein, transformants #5 and #6 have been analysed to determine the fatty acid composition of their seed oil. Several non-expressing transformants and a non-transformed control were similarly analyzed. Seed fatty acid composition is determined by the acid methanolysis method according essentially as described by Browse *et al.* (*Anal. Biochem.* (1986) *152*:141-145). Briefly, 100 seeds of each sample are treated with 1 ml of 5% H_2SO_4 in MeOH and heated in a 90°C water bath for two hours to convert the fatty acids to fatty acid methyl esters (FAMEs). An internal standard (C17:0) is added to each sample (250ml of a 1mg/ml solution in tolulene) prior to the heating step. The samples are allowed to cool, after which 1 ml 0.9% NaCl in H_20 is added to aid in phase separation. Hexane (250ml to each vial) is added to

extract the FAMEs, and the samples are then vortexed and centrifuged to separate the phases. The hexane layer is removed and transferred to a GC autosampler for injected on the GC. A useful GC temperature program for these analyses is as follows: 200°C for zero minutes, followed by a 5 degrees per minute temperature ramp to a final temperature of 250°C, which is held for 6 minutes. Data is reported as % of total fatty acids in Table I below.

Seeds from transformant #5 contain 3.95% C16:0, and seeds from #6 have a 4.59% C16:0. Seeds from the non-expressing transformants and the non-transformed control had C16:0 percentages ranging from 5.85 to 6.63%. Total saturated fatty acids in seeds from #5 were 9.74%, compared to 12.47% total saturated fatty acids for seeds from the non-transformed control and a range of 11.57%-13.33% total saturated fatty acids for seeds from the non-expressing transformants. The total saturated fatty acid level in transformant #6 is 10.64%.

TABLE I

SAMPLE:	12:0	14:0	16:0	16:1	18:0	18:1	18:2	18:3	20:0	20:1	20:2	22:0	22:1	22:2	24:0	SATS
	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%
1	0.03	0.08	6.22	0.24	3.11	18.10	25.36	18.56	2.44	21.16	1.81	0.34	2.21	0.10	0.23	12.44
3	0.08	0.09	6.08	0.24	3.02	18.62	25.21	18.75	2.26	21.21	1.70	0.31	2.09	0.12	0.21	12.06
5	0.09	0.07	3.95	0.20	2.77	17.68	27.99	18.82	2.30	20.58	2.13	0.35	2.69	0.16	0.21	9.74
6	0.01	0.07	4.59	0.18	3.15	20.95	25.30	17.71	2.28	21.33	1.75	0.32	2.04	0.09	0.22	10.64
9	0.01	0.08	5.85	0.25	2.89	19.24	25.98	17.46	2.23	21.43	1.80	0.33	2.14	0.14	0.19	11.57
10	0.11	0.12	6.63	0.33	3.14	16.48	27.66	17.07	2.71	20.59	2.16	0.38	2.24	0.14	0.24	13.33
11	0.07	0.08	6.01	0.24	3.04	19.43	24.93	17.86	2.36	21.47	1.81	0.32	2.07	0.09	0.21	12.10
12	0.01	0.08	5.91	0.21	3.09	19.98	24.28	18.84	2.23	21.16	1.59	0.33	2.02	0.09	0.18	11.83
15	0.01	0.07	5.88	0.20	3.22	20.85	24.05	18.72	2.30	20.83	1.59	0.30	1.75	0.06	0.16	11.94
CONTROL:	0.01	0.09	6.33	0.28	3.12	18.15	25.77	19.37	2.35	19.85	2.00	0.35	2.00	0.11	0.21	12.47

The above results demonstrate the ability to use synthase DNA sequences in plant genetic engineering methods for production of transgenic plants having modified seed oil compositions.

All publications and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by referenced to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claim.

What is claimed is:

1. A recombinant DNA construct comprising, in the 5' to 3' direction of transcription, a promoter functional in a plant cell, a ß-ketoacyl-ACP synthase protein encoding sequence and a transcriptional termination regulatory region 3' to said synthase protein encoding sequence, wherein said ß-ketoacyl-ACP synthase protein is from a non-plant source.

2. The construct of Claim 1, wherein said synthase protein encoding sequence is oriented for transcription of a sense sequence.

3. The construct of Claim 2, wherein said construct comprises, immediately 5' to said synthase protein encoding sequence, a plant sequence encoding a transit peptide.

4. The construct of Claim 3, wherein said synthase protein encoding sequence is the *E. coli fab*H gene sequence.

5. The construct of Claim 3, wherein said transit peptide encoding sequence is from an acyl-carrier protein gene.

6. The construct of Claim 1 wherein said promoter is from a gene that is preferentially expressed in plant seed tissue.

7. A plant cell comprising a construct according to Claim 1.

8. The plant cell of Claim 7 further comprising a second recombinant DNA construct providing for transcription in said plant cell of a substantial portion of a sequence encoding a protein associated with lipid synthesis.

9. The cell of Claim 8 wherein said protein is a desaturase or a thioesterase.

10. The cell of Claim 7 or 8 wherein said plant cell is a *Brassica* plant cell.

11. A recombinant DNA construct comprising, in the 5' to 3' direction of transcription, a promoter functional in a plant cell, a plant transit peptide encoding sequence, a mature *R. communis* ß-ketoacyl-ACP synthase factor A encoding sequence and a transcriptional termination

regulatory region 3' to said synthase protein encoding sequence, wherein said transit peptide encoding sequence is not naturally associated with said *R. communis* ß-ketoacyl-ACP synthase factor A encoding sequence.

12. The construct of Claim 11, wherein said plant transit peptide encoding sequence is from a gene encoding a *Brassica* ß-ketoacyl-ACP synthase protein.

13. The construct of Claim 12, wherein said *Brassica* synthase protein is synthase factor A.

14. The construct of Claim 11, wherein said plant transit peptide encoding sequence is from a gene encoding a stearoyl-ACP desaturase protein.

15. A plant cell comprising a construct according to Claim 11.

16. The cell of Claim 15 wherein said plant cell is a *Brassica* plant cell.

17. A transgenic plant cell comprising a non-plant ß-ketoacyl-ACP synthase protein expressed from a recombinant DNA sequence.

18. The cell of Claim 17 wherein said non-plant synthase protein is the ß-ketoacyl-ACP synthase protein encoded by *E. coli fab*H.

19. A method of producing a non-plant ß-ketoacyl-ACP synthase protein in a plant cell or progeny thereof comprising

growing a plant cell or progeny thereof comprising a construct according to Claim 1, under conditions which will permit the production of said ß-ketoacyl-ACP synthase protein.

20. A plant cell comprising a non-plant ß-ketoacyl-ACP synthase protein produced according to Claim 19.

21. A plant cell of Claim 20 wherein said construct is integrated into the genome of said plant cell.

22. A method of modifying the fatty acid composition in a plant cell comprising:

growing a plant cell having integrated in its genome a DNA construct, said construct comprising in the 5' to 3' direction of transcription, a promoter functional in said plant cell, a non-plant ß-ketoacyl-ACP synthase

protein encoding sequence, and a transcriptional termination region functional in said plant cell, under conditions which will permit the expression of said synthase protein encoding sequence.

23. The method of Claim 22 wherein said synthase protein encoding sequence is a sense sequence.

24. The method of Claim 23 wherein said synthase protein is the ß-ketoacyl-ACP synthase protein encoded by *E. coli fab*H.

25. A method of modifying the fatty acid composition in a plant cell comprising:

growing a plant cell having integrated in its genome a DNA construct according to Claim 11, under conditions which will permit the transcription of said synthase protein encoding sequence.

26. The method of Claim 22 or 25 wherein said plant cell is an oilseed plant seed cell.

27. A plant cell having a modified free fatty acid composition produced according to the method of Claim 26.

28. A plant seed having a modified fatty acid composition as compared to a seed of said plant having a native fatty acid composition, produced according to a method comprising:

growing a plant to seed, wherein said plant has integrated in its genome a recombinant DNA sequence comprising a non-plant ß-ketoacyl-ACP synthase protein encoding sequence under the transcriptional control of regulatory elements functional in seed during lipid accumulation, under conditions which will promote the activity of said regulatory elements, and

harvesting said seed.

29. The seed of Claim 28 wherein said synthase protein is the ß-ketoacyl-ACP synthase protein encoded by *E. coli fab*H.

30. A plant seed having a modified fatty acid composition as compared to a seed of said plant having a native fatty acid composition, produced according to a method comprising:

growing a plant to seed, wherein said plant has integrated in its genome a recombinant DNA construct according to Claim 11, under conditions which will promote the activity of said regulatory elements, and

harvesting said seed.

31. The seed of Claim 28 or 30 wherein said plant is an oilseed plant.

32. The seed of Claim 31 wherein said plant is *Brassica*.

33. A method of modifying the fatty acid composition of triglycerides produced from an oilseed crop plant comprising:

growing a plant cell having integrated in its genome a DNA construct, said construct comprising in the 5' to 3' direction of transcription, a promoter functional in said plant cell and a non-plant ß-ketoacyl-ACP synthase protein encoding sequence, under conditions which will permit the transcription of said synthase protein encoding sequence.

34. The method of Claim 33 wherein said synthase protein encoding sequence is a sense sequence.

35. The method of Claim 34 wherein said synthase protein is the ß-ketoacyl-ACP synthase protein encoded by *E. coli fab*H.

36. A method of modifying the fatty acid composition of triglycerides produced from an oilseed crop plant comprising:

growing a plant cell having integrated in its genome a DNA construct according to Claim 11, under conditions which will permit the transcription of said synthase protein encoding sequence.

37. The method of Claim 33 or 36 wherein said plant cell is a seed cell.

38. A plant cell having a modified fatty acid composition of triglycerides produced according to the method of Claim 37.

39. The method of Claim 37 wherein said oilseed crop plant is selected from the group consisting of rapeseed,

sunflower, safflower, cotton, cuphea, soybean, peanut, coconut, oil palm and corn.

40. A plant seed oil separated from a seed produced according to Claim 31.

41. The oil of Claim 40, wherein said plant is *Brassica*.

```
  1 GGCTTCTCCCAATTCATCGTTTGGATCGCTACCACTTCCGCCACCACCCCACCACCATGCAAGCCCTGC   69
                                                        METGlnAlaLeuG

     PstI
     |
 70 AGTCCCCGTCTCTCCGACCATCCCCTCTAACCCCGCTCCATAAAAATACTCACAATGCAGCAAAACGCC  138
    lnSerProSerLeuArgProSerProLeuThrProLeuHisLysAsnThrHisAsnAlaAlaLysArgP

139 CAACTAAAAAGGTCTCCTTTATCACCGCATCATCAACAAATAACAACACGACGATTTCAGCTCCAAAGC  207
    roThrLysLysValSerPheIleThrAlaSerSerThrAsnAsnAsnThrThrIleSerAlaProLysA

208 GAGAGAAAGACCCCAGAAAAAGGGTAGTCATAACTGGTACGGGTTTGGTATCTGTGTTTGGGAATGATG  276
    rgGluLysAspProArgLysArgValValIleThrGlyThrGlyLeuValSerValPheGlyAsnAspV

277 TCGATACTTACTACGATAAATTGCTTGCTGGAGAAAGTGGGATCGGACTTATTGATAGGTTCGATGCGT  345
    alAspThrTyrTyrAspLysLeuLeuAlaGlyGluSerGlyIleGlyLeuIleAspArgPheAspAlaS

346 CTAAGTTTCCTACTAGATTTGGTGGACAGATCAGGGGGTTTAATTCACTTGGTTATATTGATGGGAAAA  414
    erLysPheProThrArgPheGlyGlyGlnIleArgGlyPheAsnSerLeuGlyTyrIleAspGlyLysA
```

FIG. 1A

```
                                                                   SphI
                                                                   |
415 ATGATAGAAGGCTTGATGATTGTTTGAGGTATTGCATTGTTGCTGGTAAAAAGCTCTTGAGCATGCTG 483
    snAspArgArgLeuAspAspCysLeuArgTyrCysIleValAlaGlyLysLysAlaLeuGluHisAlaA

484 ATCTTGGTGGTGATAAGTTGTCTAAGATTGATAAAGAGCGAGCTGGTGTGCTTGTTGGAACAGGGATGG 552
    spLeuGlyGlyAspLysLeuSerLysIleAspLysGluArgAlaGlyValLeuValGlyThrGlyMETG

553 GTGGTCTTACAGTCTTTTCAGATGGTGTTCAGGCCCTAATTGAAAAAGGACACAGGAAAATTACCCCAT 621
    lyGlyLeuThrValPheSerAspGlyValGlnAlaLeuIleGluLysGlyHisArgLysIleThrProP

622 TCTTTATTCCTTATGCTATAACAAACATGGGATCTGCCTTGTTAGCTATTGAACTTGGTCTCATGGGTC 690
    hePheIleProTyrAlaIleThrAsnMETGlySerAlaLeuLeuAlaIleGluLeuGlyLeuMETGlyP

691 CTAATTATTCAATTTCAACTGCTTGTGCTACCTCCAATTATTGCTTCTATGCTGCTGCCAATCATATTC 759
    roAsnTyrSerIleSerThrAlaCysAlaThrSerAsnTyrCysPheTyrAlaAlaAlaAsnHisIleA

760 GCAGAGGTGAGGCTGAATTGATGATTGCTGGTGGAACTGAAGCCGCCATCATTCCAATCGGTTTGGGAG 828
    rgArgGlyGluAlaGluLeuMETIleAlaGlyGlyThrGluAlaAlaIleIleProIleGlyLeuGlyG
```

FIG. 1B

```
                                                                                       NcoI
                                                                                        |
829  GTTTTGTAGCATGTAGGGCCTTATCACAAAGGAATGATGATCCACAAACTGCCTCAAGGCCATGGGACA   897
     lyPheValAlaCysArgAlaLeuSerGlnArgAsnAspAspProGlnThrAlaSerArgProTrpAspL

898  AAGATCGAGATGGCTTTGTTATGGGTGAAGGTGCTGGAGTGTTGGTAATGGAGAGTTTGGAACATGCAA   966
     ysAspArgAspGlyPheValMETGlyGluGlyAlaGlyValLeuValMETGluSerLeuGluHisAlaM

                                 ScaI
                                  |
967  TGAAAAGGGGTGCACCAATAATTGCTGAGTACTTGGGAGGTGCTGTTAATTGTGATGCTTATCACATGA   1035
     ETLysArgGlyAlaProIleIleAlaGluTyrLeuGlyGlyAlaValAsnCysAspAlaTyrHisMETT
                                 997

1036 CTGATCCAAGGGCTGATGGACTTGGGGTCTCTTCCTGCATTGAGAGAAGTCTTGAAGATGCCGGTGTGT   1104
     hrAspProArgAlaAspGlyLeuGlyValSerSerCysIleGluArgSerLeuGluAspAlaGlyValS

                HpaI
                 |
1105 CACCTGAGGAGGTTAACTATATAAATGCACATGCAACTTCCACTCTTGCTGGTGACCTTNCTGAGATAA   1173
     erProGluGluValAsnTyrIleAsnAlaHisAlaThrSerThrLeuAlaGlyAspLeu   GluIleA
                1119
```

FIG. 1C

```
1174 ATGCTATTAAAAAGTATTCAAGAATACGTCTGACATCAAAATCAATGCAACCAAGTCTATGATAGGAC 1242
     snAlaIleLysLysValPheLysAsnThrSerAspIleLysIleAsnAlaThrLysSerMETIleGlyH

1243 ATTGCCTTGGTGCTGCTGGAGGTCTGGAAGCAATTGCCTGTGTGAAGGCCATTACCACAGGATGGTTGC 1311
     isCysLeuGlyAlaAlaGlyGlyLeuGluAlaIleAlaCysValLysAlaIleThrThrGlyTrpLeuH

1312 ATCCTACAATTAATCAATTTAACCCAGAGCCATCAGTTGAATTTGACACTGTTGCCAATAAGAAGCAGC 1380
     isProThrIleAsnGlnPheAsnProGluProSerValGluPheAspThrValAlaAsnLysLysGlnG

1381 AGCACGAAGTGAATGTTGCCATTTCAAATTCCTTTGGATTCGGTGGACACAACTCTGTGGTAGCCTTTT 1449
     lnHisGluValAsnValAlaIleSerAsnSerPheGlyPheGlyGlyHisAsnSerValValAlaPheS

                                                                   NcoI
1450 CTGCATTTAAACCCTGAGAGCATGGTTTTCTTCTGCATTCGGGCCGCGGTCATTTACATTTACCATGGC 1518
     erAlaPheLysPro .

1519 CTGCATTTCTTGTAGGAACCACTGGAGAGTTGCTTGCTTATAGACAGAGTCATCGACATCACTTCCCCC 1587

1588 TTTTAGCTTTTTGAGCTGCTGATAGTAGTCAGTTTCTCATTTCAGTATCAAGTCTATCTTAAGAAGGTC 1656

1657 TTGCTTATTTTTCTTT 1672
```

FIG. 1D

```
1   GGCTTCTCCCAATTCATCGTTGTTATCGCTACCACTTCCGCCACCACCCCACCACCATGCAAGCCCTGC   69
     LeuLeuProIleHisArgCysTyrArgTyrHisPheArgHisHisProThrThrMETGlnAlaLeuG

70  AGTCCCCGTCTCTCCGACCATCCCCTCTAACCCCGCTCCATAAAAATACTCACAATGCAGCAAAACGCC  138
    lnSerProSerLeuArgProSerProLeuThrProLeuHisLysAsnThrHisAsnAlaAlaLysArgP

139 CAACTAAAAAGGTCTCCTTTATCACCGCATCATCAACAAATAACAACACGACGATTTCAGCTCCAAAGC  207
    roThrLysLysValSerPheIleThrAlaSerSerThrAsnAsnAsnThrThrIleSerAlaProLysA

208 GAGAGAAAGACCCCAGAAAAAGGGTAGTCATAACTGGTACGGGTTTGGTATCTGTGTTTGGGAATGATG  276
    rgGluLysAspProArgLysArgValValIleThrGlyThrGlyLeuValSerValPheGlyAsnAspV

277 TCGATACTTACTACGATAAATTGCTTGCTGGAGAAAGTGGGATCGGACTTATTGATAGGTTCGATGCGT  345
    alAspThrTyrTyrAspLysLeuLeuAlaGlyGluSerGlyIleGlyLeuIleAspArgPheAspAlaS

346 CTAAGTTTCCTACTAGATTTGGTGGACAGATCAGGGGGTTTAATTCACAAGGTTATATTGATGGGAAAA  414
    erLysPheProThrArgPheGlyGlyGlnIleArgGlyPheAsnSerGlnGlyTyrIleAspGlyLysA

415 ATGATAGAAGGCTTGATGATTGTTTGAGGTATTGCATTGTTGCTGGTAAAAAGCTCTTGAGCATGCTG  483
    snAspArgArgLeuAspAspCysLeuArgTyrCysIleValAlaGlyLysLysAlaLeuGluHisAlaA
```

FIG. 2A

```
484 ATCTTGGTGGTGATAAGTTGTCTAAGATTGATAAAGAGCGAGCTGGTGTGCTTGTTGGAACAGGGATGG 552
    spLeuGlyGlyAspLysLeuSerLysIleAspLysGluArgAlaGlyValLeuValGlyThrGlyMETG

553 GTGGTCTTACAGTCTTTTCAGATGGTGTTCAGGCCCTAATTGAAAAAGGACACAGGAAAATTACCCCAT 621
    lyGlyLeuThrValPheSerAspGlyValGlnAlaLeuIleGluLysGlyHisArgLysIleThrProP

622 TCTTTATTCCTTATGCTATAACAAACATGGGATCTGCCTTGTTAGCTATTGAACTTGGTCTCATGGGTC 690
    hePheIleProTyrAlaIleThrAsnMETGlySerAlaLeuLeuAlaIleGluLeuGlyLeuMETGlyP

691 CTAATTATTCAATTTCAACTGCTTGTGCTACCTCCAATTATTGCTTCTATGCTGCTGCCAATCATATTC 759
    roAsnTyrSerIleSerThrAlaCysAlaThrSerAsnTyrCysPheTyrAlaAlaAlaAsnHisIleA

760 GCAGAGGTGAGGCTGAATTGATGATTGCTGGTGGAACTGAAGCCGCCATCATTCCAATCGGTTTGGGAG 828
    rgArgGlyGluAlaGluLeuMETIleAlaGlyGlyThrGluAlaAlaIleIleProIleGlyLeuGlyG

829 GTTTTGTAGCATGTAGGGCCTTATCACAAAGGAATGATGATCCACAAACTGCCTCAAGGCCATGGGACA 897
    lyPheValAlaCysArgAlaLeuSerGlnArgAsnAspAspProGlnThrAlaSerArgProTrpAspL

898 AAGATCGAGATGGCTTTGTTATGGGTGAAGGTGCTGGAGTGTTGGTAATGGAGAGTTTGGAACATGCAA 966
    ysAspArgAspGlyPheValMETGlyGluGlyAlaGlyValLeuValMETGluSerLeuGluHisAlaM
```

FIG. 2B

```
967  TGAAAAGGGGTGCACCAATAATTGCTGAGTACTTGGGAGGTGCTGTTAATTGTGATGCTTATCACATGA 1035
     ETLysArgGlyAlaProIleIleAlaGluTyrLeuGlyGlyAlaValAsnCysAspAlaTyrHisMETT

1036 CTGATCCAAGGGCTGATGGACTTGGGGTCTCTTCCTGCATTGAGAGAAGTCTTGAAGATGCCGGTGTGT 1104
     hrAspProArgAlaAspGlyLeuGlyValSerSerCysIleGluArgSerLeuGluAspAlaGlyValS

1105 CACCTGAGGAGGTTAACTATATAAATGCACATGCAACTTCCACTCTTGCTGGTGACCTTGCTGAGATAA 1173
     erProGluGluValAsnTyrIleAsnAlaHisAlaThrSerThrLeuAlaGlyAspLeuAlaGluIleA

1174 ATGCTATTAAAAAAGTATTCAAGAATACGTCTGACATCAAAATCAATGCAACCAAGTCTATGATAGGAC 1242
     snAlaIleLysLysValPheLysAsnThrSerAspIleLysIleAsnAlaThrLysSerMETIleGlyH

1243 ATTGCCTTGGTGCTGCTGGAGGTCTGGAAGCAATTGCCTGTGTGAAGGCCATTACCACAGGATGGTTGC 1311
     isCysLeuGlyAlaAlaGlyGlyLeuGluAlaIleAlaCysValLysAlaIleThrThrGlyTrpLeuH

1312 ATCCTACAATTAATCAATTTAACCCAGAGCCATCAGTTGAATTTGACACTGTTGCCAATAAGAAGCAGC 1380
     isProThrIleAsnGlnPheAsnProGluProSerValGluPheAspThrValAlaAsnLysLysGlnG

1381 AGCACGAAGTGAATGTTGCCATTTCAAATTCCTTTGGATTCGGTGGACACAACTCTGTGGTAGCCTTTT 1449
     lnHisGluValAsnValAlaIleSerAsnSerPheGlyPheGlyGlyHisAsnSerValValAlaPheS
```

FIG. 2C

1450 CTGCATTTAAACCCTGAGAGCATGGCCTTCTTCTGCATTCGGGCCGCGGTCATTTACATTTACCATGGC 1518
erAlaPheLysPro

1519 CTGCATTTCTTGTAGGAACCACTGGAGAGTTGCTTGCTTATAGACAGAGTCATCGACATCACTTCCCCC 1587

1588 TTTTAGCTTTTTGAGCTGCTGATAGTAGTCAGTTTCTCATTTCAGTATCAAGTCTATCTTAAGAAGGTC 1656

1657 TTGCTTAATTTTTCTTTTCAAATTACCATTTCATTGTCATTTTCCTTGGAACTTTTAGCTTAAGATCTG 1725

1726 CTGTGATCATGTGGTTTTGATTTCAAATTAATTATGTAGCGGATACGAACAAGCAATCATAAAAAGTCT 1794

1795 TTTTGAATTATGTAATTACGATAACTGTTATTTTCTTTTTCAAAAAAAAAA 1845

FIG. 2D

```
  1 CCCCGTGGCGGCGTGCATGTCGGTCACGTGCTCAAAGGAGAACAGACACGCGTTCTTCTCTTCATCGAC  69
     ProValAlaAlaCysMETSerValThrCysSerLysGluAsnArgHisAlaPhePheSerSerSerTh

 70 ACCGGGCACCACCAGCAGTCACAGTCGTACAAGAAGGAGGCCTAAATATAATAGTATCAGCACCCCTGC  138
    rProGlyThrThrSerSerHisSerArgThrArgArgArgProLysTyrAsnSerIleSerThrProAl

139 CTCTCAATCTTTCTTTAATTCTTTATCATCTTCTGGATCGAGTTTTCAACAATTAATGTCTTCTTGCTT  207
    aSerGlnSerPhePheAsnSerLeuSerSerSerGlySerSerPheGlnGlnLeuMETSerSerCysLe

208 GGCCTTCGAGCCTTGTAGTCATTACTACAGCTCTAATGGCCTCTTTCCTAACACTCCTCTTCTTCCTAA  276
    uAlaPheGluProCysSerHisTyrTyrSerSerAsnGlyLeuPheProAsnThrProLeuLeuProLy

277 GCGCCATCCTAGACTTCATCATCGCCTTCCTCGTTCTGGGGAAGCAATGGCAGTGGCTGTGCAACCTGA  345
    sArgHisProArgLeuHisHisArgLeuProArgSerGlyGluAlaMETAlaValAlaValGlnProGl

346 AAAGGAGGTTGCAACAAATAAGAAACCTCTTATGAAGCAAAGGAGAGTAGTTGTTACTGGGATGGGTGT  414
    uLysGluValAlaThrAsnLysLysProLeuMETLysGlnArgArgValValValThrGlyMETGlyVa

415 TGTTTCACCCCTTGGTCATGATATAGACGTCTATTACAATAATCTTCTTGACGGTTCTAGTGGTATTAG  483
    lValSerProLeuGlyHisAspIleAspValTyrTyrAsnAsnLeuLeuAspGlySerSerGlyIleSe
```

FIG. 3A

```
484 TCAGATTGATTCCTTTGACTGTGCCCAATTTCCTACGAGGATTGCTGGAGAGATCAAGTCTTTCTCAAC 552
    rGlnIleAspSerPheAspCysAlaGlnPheProThrArgIleAlaGlyGluIleLysSerPheSerTh

553 TGATGGATGGGTTGCACCAAAACTTTCCAAGAGAATGGATAAATTCATGCTTTACATGCTTACTGCTGG 621
    rAspGlyTrpValAlaProLysLeuSerLysArgMETAspLysPheMETLeuTyrMETLeuThrAlaGl

622 CAAAAAAGCCTTGGCAGATGGTGGTATTACAGAGGATATGATGGATGAATTGGATAAAGCTAGATGTGG 690
    yLysLysAlaLeuAlaAspGlyGlyIleThrGluAspMETMETAspGluLeuAspLysAlaArgCysGl

691 AGTTTTAATTGGTTCTGCAATGGGTGGCATGAAGGTTTTCAATGATGCAATTGAAGCATTAAGGATCTC 759
    yValLeuIleGlySerAlaMETGlyGlyMETLysValPheAsnAspAlaIleGluAlaLeuArgIleSe

760 GTATAGGAAGATGAATCCTTTCTGCGTACCTTTTGCGACTACAAATATGGGCTCTGCCATGCTTGCAAT 828
    rTyrArgLysMETAsnProPheCysValProPheAlaThrThrAsnMETGlySerAlaMETLeuAlaME

829 GGACCTTGGTTGGATGGGGCCAAACTATTCAATATCTACTGCTTGTGCTACTAGCAATTTTTGTATATT 897
    TAspLeuGlyTrpMETGlyProAsnTyrSerIleSerThrAlaCysAlaThrSerAsnPheCysIleLe

898 GAATGCCGCAAACCACATCATTAGAGGCGAAGCTGATATTATGCTTTGTGGTGGCTCAGATGCAGCAAT 966
    uAsnAlaAlaAsnHisIleIleArgGlyGluAlaAspIleMETLeuCysGlyGlySerAspAlaAlaIl
```

FIG. 3B

```
967  TATACCTATTGGCTTGGGAGGTTTTGTGGCATGCAGAGCGCTCTCACAGAGGAATGATGATCCTACAAA 1035
     eIleProIleGlyLeuGlyGlyPheValAlaCysArgAlaLeuSerGlnArgAsnAspAspProThrLy

1036 AGCTTCACGACCTTGGGATATGAATCGGGATGGATTTGTGATGGGGGAAGGAGCTGGTGTTCTTCTTTT 1104
     sAlaSerArgProTrpAspMETAsnArgAspGlyPheValMETGlyGluGlyAlaGlyValLeuLeuLe

1105 AGAAGAACTAGAACATGCTAAGAAAAGAGGTGCAAATATTTATGCGGAATTTCTTGGAGGAAGCTTTAC 1173
     uGluGluLeuGluHisAlaLysLysArgGlyAlaAsnIleTyrAlaGluPheLeuGlyGlySerPheTh

1174 ATGTGATGCTTATCACATGACTGAACCGCGTCCAGATGGAGTTGGTGTCATTCTCTGTATAGAAAAGGC 1242
     rCysAspAlaTyrHisMETThrGluProArgProAspGlyValGlyValIleLeuCysIleGluLysAl

1243 ATTAGCGCGATCTGGTGTATCCAAGGAGGAAGTAAACTACATAAATGCACATGCTACGTCTACCCCAGC 1311
     aLeuAlaArgSerGlyValSerLysGluGluValAsnTyrIleAsnAlaHisAlaThrSerThrProAl

1312 TGGAGACCTTAAAGAATATGAAGCTCTTATGCGCTGTTTCAGCCAAAATCCTGATTTGAGAGTGAACTC 1380
     aGlyAspLeuLysGluTyrGluAlaLeuMETArgCysPheSerGlnAsnProAspLeuArgValAsnSe

1381 TACGAAGTCTATGATTGGCCATTTACTAGGAGCAGCTGGTGCTGTGGAAGCTATAGCAACAATACAGGC 1449
     rThrLysSerMETIleGlyHisLeuLeuGlyAlaAlaGlyAlaValGluAlaIleAlaThrIleGlnAl
```

FIG. 3C

```
1450 GATACGGACAGGATGGGTTCATCCAAACATCAACCTGGAAAACCCAGAAGAAGGCGTGGACACAAAGGT 1518
     aIleArgThrGlyTrpValHisProAsnIleAsnLeuGluAsnProGluGluGlyValAspThrLysVa

1519 GCTGGTTGGCCCAAAGAAGGAGAGATTGGACATTAAGGTTGCTCTGTCCAACTCTTTTGGGTTCGGTGG 1587
     lLeuValGlyProLysLysGluArgLeuAspIleLysValAlaLeuSerAsnSerPheGlyPheGlyGl

1588 GCACAACTCATCGATCATTTTTGCTCCGTACAAGTGAAATAAGGGGTACTTCAACTTTGGTGTATTAAC 1656
     yHisAsnSerSerIleIlePheAlaProTyrLys

1657 GTGAAAGATGATCTAAAATGGAACAAGATTAGATAACTCTATGGGTAGGGAAAGGAGAATATGCCGAGT 1725

1726 TCACAGAGAGGAAACTTCCCGTGAAGATTCCTGTGCCTTCTACCATTTTCAGTATTCTCTCCGCATCAT 1794

1795 TGTGGCTTGATCCATGTTGATCCATCGAATACCAGTAACAGTGGCCTTATTTAATTTTTGTTCCATGTA 1863

1864 TAAGCAGACGGCTGATCGTTGCTTTAACAGTCAATTGTAATGAATTTTTGAGCTGGACAGTTGGCTAGG 1932

1933 TTACACTAATGTAATGGTGGTTTTATGAGCAAAAAAA 1969
```

FIG. 3D

```
  1 ATGCGAGACAGCCCACGAGAAGACGCTCATTCATCTCCGCGTCGTCCTCCGCCGTCTCCGCCCCCAAAC  69
     AlaArgGlnProThrArgArgArgSerPheIleSerAlaSerSerSerAlaValSerAlaProLysA

 70 GCGAAACAGACCCGAAGAAACGGGTCGTAATCACCGGAATGGCCTCGTCTCCGTCTTCGGAAACGACG 138
    rgGluThrAspProLysLysArgValValIleThrGlyMETGlyLeuValSerValPheGlyAsnAspV

139 TCGACGCTTACTACGAGAAGCTGCTCTCCGGCGAGAGTGGAATCAGCTTGATTGATCGGTTCGACGCCT 207
    alAspAlaTyrTyrGluLysLeuLeuSerGlyGluSerGlyIleSerLeuIleAspArgPheAspAlaS

208 CCAAGTTCCCGACCCGATTCGGTGGACAGATCCGTGGGTTCAGCTCAGAGGGTTACATCGATGGGAAGA 276
    erLysPheProThrArgPheGlyGlyGlnIleArgGlyPheSerSerGluGlyTyrIleAspGlyLysA

277 ATGAGCGGAGGCTTGATGATTGCTTGAAGTACTGCATTGTCGCTGGGAAGAAGGCTCTTGAAAGTGCGA 345
    snGluArgArgLeuAspAspCysLeuLysTyrCysIleValAlaGlyLysLysAlaLeuGluSerAlaA

346 ATCTTGGTGGTGATAAGCTTAACACGATTGATAAGCAGAAAGCTGGAGTACTAGTTGGGACTGGTATGG 414
    snLeuGlyGlyAspLysLeuAsnThrIleAspLysGlnLysAlaGlyValLeuValGlyThrGlyMETG

415 GTGGCTTGACTGTGTTTTCAGACGGTGTTCAAGCTCTTATTGAGAAAGGTCACAGGAGGATTTCTCCTT 483
    lyGlyLeuThrValPheSerAspGlyValGlnAlaLeuIleGluLysGlyHisArgArgIleSerProP
```

FIG. 4A

```
484 TCTTTATTCCTTATGCTATTACAAACATGGGTTCTGCTTTGTTGGCGATTGATCTTGGTCTTATGGGTC 552
    hePheIleProTyrAlaIleThrAsnMETGlySerAlaLeuLeuAlaIleAspLeuGlyLeuMETGlyP

553 CTAACTACTCGATCTCGACGGCTTGTGCCACTTCTAACTACTGCTTTTACGCTGCTGCGAATCACATTC 621
    roAsnTyrSerIleSerThrAlaCysAlaThrSerAsnTyrCysPheTyrAlaAlaAlaAsnHisIleA

622 GACGTGGTGAAGCTGATATGATGATAGCTGGTGGAACCGAGGCTGCTATTATTCCTATTGGTTTGGGAG 690
    rgArgGlyGluAlaAspMETMETIleAlaGlyGlyThrGluAlaAlaIleIleProIleGlyLeuGlyG

691 GTTTTGTTGCTTGTAGGGCGCTTTCACAGAGAAATGATGATCCTCAGACGGCTTCAAGGCCGTGGGATA 759
    lyPheValAlaCysArgAlaLeuSerGlnArgAsnAspAspProGlnThrAlaSerArgProTrpAspL

760 AACAGAGAGATGGGTTTGTCATGGGTGAAGGAGCTGGTGTTCTGGTGATGGAAAGCTTGGAACATGCGA 828
    ysGlnArgAspGlyPheValMETGlyGluGlyAlaGlyValLeuValMETGluSerLeuGluHisAlaM

829 TGAAACGTGGTGCTCCAATTGTAGCAGAGTATCTTGGAGGCGCTGTTAACTGCGATGCTCATCATATGA 897
    ETLysArgGlyAlaProIleValAlaGluTyrLeuGlyGlyAlaValAsnCysAspAlaHisHisMETT

898 CTGATCCAAGAGCTGATGGGCTTGGTGTGTCTTCATGCATTGAGAGCTGCCTTGAAGATGCTGGTGTAT 966
    hrAspProArgAlaAspGlyLeuGlyValSerSerCysIleGluSerCysLeuGluAspAlaGlyValS
```

FIG. 4B

```
967  CACCTGAGGAGGTAAATTACATCAATGCACATGCAACTTCCACACTGGCTGGTGATCTTGCTGAGATTA 1035
     erProGluGluValAsnTyrIleAsnAlaHisAlaThrSerThrLeuAlaGlyAspLeuAlaGluIleA

1036 ATGCCATTAAAAAGGTATTCAAAAGCACTTCAGGGATCAAAATCAATGCCACCAAGTCTATGATAGGTC 1104
     snAlaIleLysLysValPheLysSerThrSerGlyIleLysIleAsnAlaThrLysSerMETIleGlyH

1105 ACTGCCTCGGTGCAGCTGGAGGTCTTGAAGCCATTGCCACCGTGAAGGCTATCAACACGGGATGGCTGC 1173
     isCysLeuGlyAlaAlaGlyGlyLeuGluAlaIleAlaThrValLysAlaIleAsnThrGlyTrpLeuH

1174 ATCCCTCTATCAACCAATTTAACCCAGAACCAGCAGTGGACTTTGATACGGTCGCAAACGAGAAGAAGC 1242
     isProSerIleAsnGlnPheAsnProGluProAlaValAspPheAspThrValAlaAsnGluLysLysG

1243 AGCATGAGGTGAATGTTGCCATATCAAACTCGTTTGGGTTCGGTGGACATAACTCAGTGGTCGCTTTCT 1311
     lnHisGluValAsnValAlaIleSerAsnSerPheGlyPheGlyGlyHisAsnSerValValAlaPheS

1312 CTGCCTTCAAACCCTGATTTCCTCAGACCCTTTAGATCCTCTGGTCCATCTGTTAGATCACCACCATCA 1380
     erAlaPheLysPro

1381 TCTTCTTCGCAGCTTCTTGGTTCACAAGTTGAGCGCTTTCTTCCTTTCAGCTTTTGTTCTTATTGGTC 1449
```

FIG. 4C

1450 ATTGTTAATTTTTGCTCAACTCTTATTGGTCATTGAGGTGTAGAGAATCCAGATTTTGCTTCTACAATC 1518

1519 TGTGTACGGAATGTTGTATCTTTAGTTCGTTTTATGTTTGCCAAATTTTATAAAC 1573

FIG. 4D

```
1   AACCACATTCGCCGTGGGGAAGCTGATATGATGATTGCTGGTGGAACCGAGGCTGCCATTATTCCTATT   69
    AsnHisIleArgArgGlyGluAlaAspMETMETIleAlaGlyGlyThrGluAlaAlaIleIleProIle

70  GGGTTGGGAGGTTTTGTTGCTTGCAGGGCGCTTTCGCAGAGGAATGATGACCCTAAAACCGCTTCGAGG   138
    GlyLeuGlyGlyPheValAlaCysArgAlaLeuSerGlnArgAsnAspAspProLysThrAlaSerArg

139 CCTTGGGATAAACAGAGAGATGGCTTTGTAATGGGTGAAGGAGCTGGTGTTCTGGTGATGGAAAGCTTG   207
    ProTrpAspLysGlnArgAspGlyPheValMETGlyGluGlyAlaGlyValLeuValMETGluSerLeu

208 GAACATGCGATGAAGCGTGGTGCGCCAATAGTAGCAGAGTATCTTGGAGGTGCTGTAAACTGTGATGCT   276
    GluHisAlaMETLysArgGlyAlaProIleValAlaGluTyrLeuGlyGlyAlaValAsnCysAspAla

277 CATCATATGACTGATCCAAGAGCTGACgGGCTTGGTGTCTCTTCATGCATTGAGAGCTGCCTTGAAGAT   345
    HisHisMETThrAspProArgAlaAspGlyLeuGlyValSerSerCysIleGluSerCysLeuGluAsp

346 GCTGGTGTTTCACCCGAGGAGGTAAATTACATCAATGCGCATGCAACTTCCACACTTGCTGGTGATCTT   414
    AlaGlyValSerProGluGluValAsnTyrIleAsnAlaHisAlaThrSerThrLeuAlaGlyAspLeu

415 GCTGAGATTAATGCCATTAAAAAGGTATTCAAGAGCACTGCTGGGATCAAAATCAATGCCACCAAGTCT   483
    AlaGluIleAsnAlaIleLysLysValPheLysSerThrAlaGlyIleLysIleAsnAlaThrLysSer
```

FIG. 5A

```
484 ATGATAGGTCACTGCCTCGGTGCAGCTGGAGGTCTTGAAGCCATTGCGACTGTGAAGGCTATCAACACT 552
    METIleGlyHisCysLeuGlyAlaAlaGlyGlyLeuGluAlaIleAlaThrValLysAlaIleAsnThr

553 GGATGGCTTCATCCCTCAATCAACCAATTTaaCCCAGAACCAGCCGTGGACTTTGACACGGTCGCAAAC 621
    GlyTrpLeuHisProSerIleAsnGlnPheAsnProGluProAlaValAspPheAspThrValAlaAsn

622 GAGAAGAAGCAGCATGAGGTGAACGTTGCTATATCAAATTCGTTTGGGTTCGGTGGACACAACTCAGTT 690
    GluLysLysGlnHisGluValAsnValAlaIleSerAsnSerPheGlyPheGlyGlyHisAsnSerVal

691 GTCGCCTTCTCTGCCTTCAAACCCTGATTCCTTCAAGACCCTTTTGTATTTTCTTCTCCAACTATTACA 759
    ValAlaPheSerAlaPheLysPro

760 TCACCACCATCATCCATCAGGCATCATCTTCCTTGAGCTTCTTGGTTCCACGAGTTTGAGCTCTTTCTT 828

829 TGGCGTTTTACGTTCCATTCAACATTGTTCTTATTGTTCATTGAGATTTCAAATTTTGCTTCTCAATCG 897

898 TAAGAAATGTTTGTATCTGTATCTGTATCTGAGTTCGTTTCATATTTGTCTAATTTATAAACAGAACCA 966

967 ATAATCTTGTAGCAATGATGTTATTCAGAGTTCTCAATCTT 1007
```

FIG. 5B

```
CCCCCCGACG CGTCCAAACA CTCAAGTGTG AGAGAGAGAT CAGATAATCT TTCTCGTTTT   60
CTCCACCTTC ATCCGAGTAT GACGATGGGT GGTGCGTCTT TATGCGATTC ACTAGTGGCT  120
GCTTGCATGT CCTCCGCCTC GCACTCAAGC GGAGACCGAC TGACTCAATT CATCTGGCCT  180
CGCCGGAGTA GACTGGTTAA CAACTGCTCG CTCCATGGAT CCCAGGCGAG TTCCCGTAAC  240
AACAATGCCT CGTCTTCCCT CTTCGAATCG AATAACACTT CCTTCAATCC AAAGCAGAGG  300
AGATTCAATC GAGCATCAAC CTCTGGGCAA GTCACTACAC TAGAGATGGA GAAGGACGCA  360
ATGGTAAACA AGCCACGCCG AGTTGTTGTC ACTGGCATGG GAGTTGAAAC ACCACTAGGT  420
CACGACCCTC ATACTTTTTA TGACAACTTG CTACAAGGCA AAAGTGGTAT AAGCCATATA  480
GAGAGTTTCG ACTGTTCTGC ATTTCCCACT AGAATCGCTG GGGAGATTAA ATCTTTTTCG  540
ACCGACGGAT TGGTTGCTCC TAAACTTTCC AAAAGGATGG ACAAGTTCAT GCTCTACCTT  600
```

FIG. 6A

CTAACCGCCG GCAAGAAGGC GTTGGAGGAT GGTGGGGTGA CTGGGGATGT GATGGCAGAG 660

TTCGACAAAT CAAGATGTGG TGTCTTGATT GGCTCAGCAA TGGGAGGCAT GAAGGTCTTT 720

TACGATGCGC TTGAAGCTTT GAAAATCTCT TACAGGAAGA TGAACCCTTT TTGTGTACCT 780

TTTGCCACCA CAAACATGGG TTCCGCTATG CTTGCCTTGG ATCTGGGATG GATGGGTCCA 840

AACTACTCTA TTTCAACCGC ATGTGCCACG GGAAACTTCT GTATTCTCAA TGCGGCAAAC 900

CACATTACCA GAGGTGAAGC TGATGTAATG CTCTGTGGTG GCTCTGACTC AGTTATTATT 960

CCAATAGGGT TGGGAGGTTT TGTTGCCTGC CGGGCTCTTT CAGAAAATAA TGATGATCCC 1020

ACCAAAGCTT CTCGTCCTTG GGATAGTAAC CGAGATGGTT TTGTTATGGG AGAGGGAGCC 1080

GGAGTTCTAC TTTTAGAAGA ACTTGAGCAT GCCAAGAGGA GCAACTATAT ACGCAGAGTT 1140

CCTTGGGGGT AGTTTCACAT GTGATGCATA CCATATAACC GAACCACGTC CTGATGGTGC 1200

FIG. 6B

```
TGGTGTCATT CTTGCTATCG AGAAAGCGGT AGCTCATGCC GGGATTTCTA AGGAAGACAT 1260
AAATTACGTG AATGCTCATG CTACCTCTAC ACCAGCTGGA GACCTTAAGG AGTACCACGC 1320
TCTTTCTCAC TGTTTTGGCC AAAATCCTGA GCTAAGAGTA AACTCAACAA AATCTATGAT 1380
TGGACACTTG CTGGGAGCTT CTGGGGCCGT GGAGGCTGTT GCAACCGTTC AGGCAATAAA 1440
GACAGGATGG GTTCATCCAA ATATCAACCT CGAGAATCCA GACAAAGCAG TGGATACAAA 1500
GCTTTTGGTG GGTCTTAAGA AGGAGAGACT GGATATCAAA GCAGCCTTGT CAAACTCTTT 1560
CGGCTTTGGT GGCCAGAACT CTAGCATAAT TTTCGCTCCT TACAAATGAA AGGCGAATAG 1620
TCCAATGCTG TGTACTCTTG TGTAACTTGC TGTAAGTGTG TACAAGAACT TCCCATGTTT 1680
TGATGCAATA TGTACGAGAA CTTCCCATGC TTTTGGTAGT GCCATGATTC AGGATTCGAT 1740
TAACTTGCAC AAAGAGTTTA AGCAACGTTG AAAAGAGAGA GAAAAAAAAA GTGATGAGGT 1800
```

FIG. 6C

```
AGCTGAGGAT TTGTCAGGAA CAACAATACT TCATTTTTCA CTTTGGTTAG GTAGACTGAA 1860
ATATTTGAGC CAACATTTCT TGTATTTTTA TTCTTTGAAA GCTTTAACCA AAGAAAAAAA 1920
AA                                                                1922
```

FIG. 6D

```
Met Thr Met Gly Gly Ala Ser Leu Cys Asp Ser Leu Val Ala Ala Cys
 1              5                   10                  15

Met Ser Ser Ala Ser His Ser Ser Gly Asp Arg Leu Thr Gln Phe Ile
            20                  25                  30

Trp Pro Arg Arg Ser Arg Leu Val Asn Asn Cys Ser Leu His Gly Ser
        35                  40                  45

Gln Ala Ser Ser Arg Asn Asn Asn Ala Ser Ser Ser Leu Phe Glu Ser
    50                  55                  60

Asn Asn Thr Ser Phe Asn Pro Lys Gln Arg Arg Phe Asn Arg Ala Ser
65                  70                  75                  80

Thr Ser Gly Gln Val Thr Thr Leu Glu Met Glu Lys Asp Ala Met Val
                85                  90                  95

Asn Lys Pro Arg Arg Val Val Val Thr Gly Met Gly Val Glu Thr Pro
            100                 105                 110

Leu Gly His Asp Pro His Thr Phe Tyr Asp Asn Leu Leu Gln Gly Lys
        115                 120                 125
```

FIG. 7A

```
Ser Gly Ile Ser His Ile Glu Ser Phe Asp Cys Ser Ala Phe Pro Thr
    130                 135                 140

Arg Ile Ala Gly Glu Ile Lys Ser Phe Ser Thr Asp Gly Leu Val Ala
145                 150                 155                 160

Pro Lys Leu Ser Lys Arg Met Asp Lys Phe Met Leu Tyr Leu Leu Thr
                165                 170                 175

Ala Gly Lys Lys Ala Leu Glu Asp Gly Gly Val Thr Gly Asp Val Met
            180                 185                 190

Ala Glu Phe Asp Lys Ser Arg Cys Gly Val Leu Ile Gly Ser Ala Met
        195                 200                 205

Gly Gly Met Lys Val Phe Tyr Asp Ala Leu Glu Ala Leu Lys Ile Ser
    210                 215                 220

Tyr Arg Lys Met Asn Pro Phe Cys Val Pro Phe Ala Thr Thr Asn Met
225                 230                 235                 240
```

FIG. 7B

```
Gly Ser Ala Met Leu Ala Leu Asp Leu Gly Trp Met Gly Pro Asn Tyr
                245                 250                 255

Ser Ile Ser Thr Ala Cys Ala Thr Gly Asn Phe Cys Ile Leu Asn Ala
            260                 265                 270

Ala Asn His Ile Thr Arg Gly Glu Ala Asp Val Met Leu Cys Gly Gly
        275                 280                 285

Ser Asp Ser Val Ile Ile Pro Ile Gly Leu Gly Gly Phe Val Ala Cys
    290                 295                 300

Arg Ala Leu Ser Glu Asn Asn Asp Asp Pro Thr Lys Ala Ser Arg Pro
305                 310                 315                 320

Trp Asp Ser Asn Arg Asp Gly Phe Val Met Gly Glu Gly Ala Gly Val
                325                 330                 335

Leu Leu Leu Glu Glu Leu Glu His Ala Lys Arg
            340                 345
```

FIG. 7C

```
Ile Tyr Ala Glu Phe Leu Gly Gly Ser Phe Thr Cys Asp Ala Tyr His>
 1               5                   10                  15

Ile Thr Glu Pro Arg Pro Asp Gly Ala Gly Val Ile Leu Ala Ile Glu>
            20                  25                  30

Lys Ala Val Ala His Ala Gly Ile Ser Lys Glu Asp Ile Asn Tyr Val>
        35                  40                  45

Asn Ala His Ala Thr Ser Thr Pro Ala Gly Asp Leu Lys Glu Tyr His>
    50                  55                  60

Ala Leu Ser His Cys Phe Gly Gln Asn Pro Glu Leu Arg Val Asn Ser>
65                  70                  75                  80

Thr Lys Ser Met Ile Gly His Leu Leu Gly Ala Ser Gly Ala Val Glu>
                85                  90                  95

Ala Val Ala Thr Val Gln Ala Ile Lys Thr Gly Trp Val His Pro Asn>
            100                 105                 110
```

FIG. 8A

```
Ile Asn Leu Glu Asn Pro Asp Lys Ala Val Asp Thr Lys Leu Leu Val>
        115                 120                 125

Gly Leu Lys Lys Glu Arg Leu Asp Ile Lys Ala Ala Leu Ser Asn Ser>
    130                 135                 140

Phe Gly Phe Gly Gly Gln Asn Ser Ser Ile Ile Phe Ala Pro Tyr Lys>
145                 150                 155                 160
```

FIG. 8B

```
  1 ATGATTACCTGAAAATAAGTATAATTTGTATTGAAATTATAAAGTGACATTTTTTGTGTAACAAATATT  69

 70 TTGTGTAACAAGAATTAAAAAAAAAAACAGAAAATACTCAGCTTTTTTAATAATAAAAAAAATTAATTG 138

139 AGTTAGAAAATTGTTGTACCAATAACAAAAGATTTATATGGAATTATAAAATCAACACACCAATAACAC 207

208 AAGACTTTTTAAAAATTTAAGAATAATATAAGCAATAACAATAGAATCTTCAAATTCTTCAAATCCTTA 276

277 AAAATCAATCTCCCACTATTAATCCCCCTTAGTTTTAGTTGGTAATGGCAACGTTTGTTGACTACCGTA 345

346 TTGTAACTTTTGTCAAATTGTCATAAATACGTGTCAAACTCTGGTAAAAAATTAGTCTGCTACATCTGT 414

415 CTTTTATTTATAAAACACAGCTGTTAATCAGAATTTGGTTTATTAAATCAACAACCTGCACGAAACTTG 483

484 TGTGAGCATATTTTGTCTGTTTCTGGTTCATGACCTTCTTCCGCATGATGGCCAAGTGTAATGGCCACT 552
```

FIG. 9A

```
                                                                        BglII
                                                                        |
553 TGCAAGAGCGTTTCTTCAACGAGATAAGTCGAACAAATATTTGTCCGTTACGACCACATATAANATCTC  621
                                                                        616

622 CCCATCTCTATATATAATACCAGCATTCACCATCATGAATACCTCAAATCCCAATCTCACAAATACTTC  690

691 AATAAAAAGACCAAAAAAAATTAAAGCAAAGAAAAGCCTTCTTGTGCACAAAAAAAAAAGAAGCCTTCT  759

760 AGGTTTTCACGACATGAAGTTCACTACTCTAATGGTCATCACATTGGTGATAATCGCCATCTCGTCTCC  828
                 METLysPheThrThrLeuMETValIleThrLeuValIleIleAlaIleSerSerPr

829 TGTTCCAATTAGAGCAACCACGGTTGAAAGTTTCGGAGAAGTGGCACAATCGTGTGTTGTGACAGAACT  897
    oValProIleArgAlaThrThrValGluSerPheGlyGluValAlaGlnSerCysValValThrGluLe

898 CGCCCCATGCTTACCAGCAATGACCACGGCAGGAGACCCGACTACAGAATGCTGCGACAAACTGGTAGA  966
    uAlaProCysLeuProAlaMETThrThrAlaGlyAspProThrThrGluCysCysAspLysLeuValGl

967 GCAGAAACCATGTCTTTGTGGTTATATTCGAAACCCAGCCTATAGTATGTATGTTACTTCTCCAAACGG 1035
    uGlnLysProCysLeuCysGlyTyrIleArgAsnProAlaTyrSerMETTyrValThrSerProAsnGl
```

FIG. 9B

```
1036 TCGCAAAGTCTTAGATTTTTGTAAGGTTCCTTTTCCTAGTTGTTAAATCTCTCAAGACATTGCTAAGAA 1104
     yArgLysValLeuAspPheCysLysValProPheProSerCys .

                                        BglII                               HindII
                                        |                                   |
1105 AAATATTATTAAAAATAAAAGAATCAAACTAGATCTGATGTAACAATGAATCATCATGTTATGGTTGAA 1173
                                        1136                                1173

1174 GCTTATATAGCTGAAGTGTTTTGATTTTATATATGTGTGTGTGTGTGTCCTGCTCAATTTTTGAAACAC 1242

1243 ACACGTTTCTCCTGATTTGGATTTAAATTATATTTTGAGTTAAAAAAAAGAAAAAGATGGAATGCTATT 1311

                                              EcoRV
                                              |
1312 TATACAAGTTGATGAAAAAGTGGAAGTACAATTTAGATATCTCCWWCACTTAAAGAATGAAACAATAAT 1380
                                              1350

                                         SalI
                                         |
1381 AGACTTCGAAACAAATGAAAAATACATAAATTGTCGACAATCAACGTCGATCGACGAGTTTATTATTAA 1449
                                         1414
```

FIG. 9C

```
1450 AAATTTGTGTGAAGGACTAGCAGTTCAACCAAATGATATTGAACATATACATCAACAAATATGATAATC 1518

1519 ATAAAAGAGAGAATGGGGGGGGGGTGTCGTTTACCAGAAACCTCTTTTTCTCAGCTCGCTAAAACCCTA 1587

1588 CCACTAGAGACCTAGCTCTGACCGTCGGCTCATCGGTGCCGGAGGTGTAACCTTTCTTTCCCATGACCC 1656

1657 GAAACCTCTCTTTCCCAACTCACGAAAACCCTACAATCAAAAACCTAGCTCCGACCATCGGCTCATCGG 1725

                                                                      ClaI
1726 TGCCGAAGGTGTAACCTTTCNCTCCCATCATAGTTTCTCGTAAATGAAAGCTAATTGGGCAATCGATTT 1794
                                                                      1789

1795 TTTAATGTTTAAACCATGCCAAGCCATTTCTTATAGGACAATTGTCAATAATAGCATCTTTTGAGTTTT 1863

1864 GTCTCAAAAGTGACACTAGAAGAAAAAAGTCACAAAAATGACATTCATTAAAAGTAAAATATCCCTAA 1932
```

FIG. 9D

```
1933 TACCTTTGGTTTAAATTAAATAAGTAAACAAAAATAAATAAAAACAAATAAAATAAAAATAAAAAATGA 2001

2002 AAAAAAGAAATTTTTTTATAGTTTCAGATTATATGTTTTCAGATTCGAAATTTTTTAAA          2060
```

FIG. 9E

```
                                            HindIII
                                            |
1   GCTCACTTGTGTGGTGGAGGAGAAAAACAGAACTCACAAAAAGCTTTGCGACTGCCAAGAACAACAACA   69
                                            42

70  ACAACAAGATCAAGAAGAAGAAGAAGAAGATCAAAAATGGCTCTTCGAATCACTCCAGTGACCTTGCAA  138
                                         METAlaLeuArgIleThrProValThrLeuGln

          EcoRV
          |                                 BglII          NcoI
          |                                 |              |
139 TCGGAGAGATATCGTTCGTTTCGTTTCCTAAGAAGGCTAATCTCAGATCTCCCAAATTCGCCATGGCC  207
    SerGluArgTyrArgSerPheSerPheProLysLysAlaAsnLeuArgSerProLysPheAlaMETAla
          149                               185            201

                         HindII
                         |
208 TCCACCCTCGGATCATCCACACCGAAGGTTGACAATGCCAAGAAGCCTTTTCAACCTCCACGAGAGGTT  276
    SerThrLeuGlySerSerThrProLysValAspAsnAlaLysLysProPheGlnProProArgGluVal
                         238
```

FIG. 10A

```
277 CATGTTCAGGTGACGCACTCCATGCCACCACAGAAGATAGAGATTTTCAAATCCATCGAGGGTTGGCT  345
    HisValGlnValThrHisSerMETProProGlnLysIleGluIlePheLysSerIleGluGlyTrpAla

346 GAGCAGAACATATTGGTTCACCTAAAGCCAGTGGAGAAATGTTGGCAAGCACAGGATTTCTTGCCGGAC  414
    GluGlnAsnIleLeuValHisLeuLysProValGluLysCysTrpGlnAlaGlnAspPheLeuProAsp

415 CCTGCATCTGAAGGATTTGATGAACAAGTCAAGGAACTAAGGGCAAGAGCAAAGGAGATTCCTGATGAT  483
    ProAlaSerGluGlyPheAspGluGlnValLysGluLeuArgAlaArgAlaLysGluIleProAspAsp

484 TACTTTGTTGTTTTGGTTGGAGATATGATTACAGAGGAAGCCCTACCTACTTACCAAACAATGCTTAAT  552
    TyrPheValValLeuValGlyAspMETIleThrGluGluAlaLeuProThrTyrGlnThrMETLeuAsn

553 ACCCTAGATGGTGTACGTGATGAGACTGGGGCTAGCCTTACGCCTTGGGCTGTCTGGACTAGGGCTTGG  621
    ThrLeuAspGlyValArgAspGluThrGlyAlaSerLeuThrProTrpAlaValTrpThrArgAlaTrp

        PvuII
        |                                                          AccI
622 ACAGCTGAAGAGAACAGGCATGGCGATCTTCTCCACACCTATCTCTACCTTTCTGGCGGGTAGACATG  690
    ThrAlaGluGluAsnArgHisGlyAspLeuLeuHisThrTyrLeuTyrLeuSerGlyArgValAspMET
        626                                                        684
```

FIG. 10B

```
                                                                BamHI
                                                                |
691 AGGCAGATACAGAAGACAATTCAGTATCTCATTGGGTCAGGAATGGATCCTCGTACCGAAAACAGCCCC 759
    ArgGlnIleGlnLysThrIleGlnTyrLeuIleGlySerGlyMETAspProArgThrGluAsnSerPro
                                                                736

760 TACCTTGGGTTCATCTACACATCGTTTCAAGAGCGTGCCACATTTGTTTCTCACGGAAACACCGCCAGG 828
    TyrLeuGlyPheIleTyrThrSerPheGlnGluArgAlaThrPheValSerHisGlyAsnThrAlaArg

       SphI
       |
829 CATGCAAAGGATCATGGGGACGTGAAACTGGCGCAAATTTGTGGTACAATCGCGTCTGACGAAAAGCGT 897
    HisAlaLysAspHisGlyAspValLysLeuAlaGlnIleCysGlyThrIleAlaSerAspGluLysArg
       833

                                                                ClaI
                                                                |
898 CACGAGACCGCTTATACAAAGATAGTCGAAAAGCTATTCGAGATCGATCCTGATGGCACCGTTCTTGCT 966
    HisGluThrAlaTyrThrLysIleValGluLysLeuPheGluIleAspProAspGlyThrValLeuAla
                                                                942
```

FIG. 10C

```
                         BglII
                         |
967  TTTGCCGACATGATGAGGAAAAAGATCTCGATGCCCGCACACTTGATGTACGATGGGCGTGATGACAAC 1035
     PheAlaAspMETMETArgLysLysIleSerMETProAlaHisLeuMETTyrAspGlyArgAspAspAsn
                         990

                                               AccI
                                               |
1036 CTCTTCGAACATTTCTCGGCGGTTGCCCAAAGACTCGGCGTCTACACCGCCAAAGACTACGCCGACATA 1104
     LeuPheGluHisPheSerAlaValAlaGlnArgLeuGlyValTyrThrAlaLysAspTyrAlaAspIle
                                               1077

1105 CTGGAATTTCTGGTCGGGCGGTGGAAAGTGGCGGATTTGACCGGCCTATCTGGTGAAGGGCGTAAAGCG 1173
     LeuGluPheLeuValGlyArgTrpLysValAlaAspLeuThrGlyLeuSerGlyGluGlyArgLysAla

                                                         SacI
                                                         |
1174 CAAGATTATGTTTGCGGGTTGCCACCAAGAATCAGAAGGCTGGAGGAGAGAGCTCAAGGGCGAGCAAAG 1242
     GlnAspTyrValCysGlyLeuProProArgIleArgArgLeuGluGluArgAlaGlnGlyArgAlaLys
                                                         1228
```

FIG. 10D

```
                                            PvuII
                                              |
1243 GAAGGACCTGTTGTTCCATTCAGCTGGATTTTCGATAGACAGGTGAAGCTGTGAAGAAAAAAAAAACGA 1311
     GluGlyProValValProPheSerTrpIlePheAspArgGlnValLysLeu
                                            1266

1312 GCAGTGAGTTCGGTTTCTGTTGGCTTATTGGGTAGAGGTTAAAACCTATTTAGATGTCTGTTTCGTGT 1380

1381 AATGTGGTTTTTTTTCTTCTAATCTTGAATCTGGTATTGTGTCGTTGAGTTCGCGTGTGTGTAAACTTG 1449

1450 TGTGGCTGTGGACATATTATAGAACTCGTTATGCCAATTTTGATGACGGTGGTTATCGTCTCCCCTGGT 1518

1519 GTTTTTTTATTGTTT 1533
```

FIG. 10E

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/10526

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :C07H 21/04; C12N 5/00, 15/00; C12P 7/64, 21/04
US CL :435/69.7, 69.8, 70.1, 134, 172.3, 240.4; 536/23.2, 23.4, 23.7, 24.1
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.7, 69.8, 70.1, 134, 172.3, 240.4; 536/23.2, 23.4, 23.7, 24.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 5,110,728 (KRIDL ET AL) 05 MAY 1992, see column 1, lines 20-42; column 9, lines 51-64; column 10, lines 61-68; column 15, lines 3-55; column 16, lines 8-12 and 48-51.	1-7, 10, 17-24, 26, 33-35, 37, 39
Y	TRENDS IN BIOTECHNOLOGY, Volume 5, No. 2, issued 1987, V. C. Knauf, "The Application of Genetic Engineering to Oilseed Crops", pages 40-47, see pages 43-45.	1-7, 10, 17-24, 26, 33-35, 37, 39
Y	CARLSBERG RESEARCH COMMUNICATIONS, Volume 53, issued 1988, S. Kauppinen et al, "β-Ketoacyl-ACP Synthase I of Escherichia coli: Nucleotide Sequence of the fabB Gene and Identification of the Cerulenin Binding Residue", pages 357-370, see page 361.	1-7, 10, 17-24, 26, 33-35, 37, 39

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be part of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
04 JANUARY 1994	07 FEB 1994

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	DAVID T. FOX D. Vargza for
Facsimile No. NOT APPLICABLE	Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/10526

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	THE JOURNAL OF BIOLOGICAL CHEMISTRY, Volume 267, No. 10, issued 05 April 1992, J. Tsay et al, "Isolation and Characterization of the β-Ketoacyl-acyl Carrier Protein Synthase III Gene (fabH) from Escherichia coli K-12", pages 6807-6814, see page 6810.	4, 18, 24, 35

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/10526

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
(Telephone Practice)
Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-7, 10, 17-24, 26, 33-35, 37 and 39

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/10526

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

I. Claims 1-7, 10, 17-24, 26, 33-35, 37 and 39, drawn to a recombinant DNA construct comprising a bacterial β-ketoacyl-ACP synthase gene, plant cells containing the construct, and a method of use, classified in Class 435, Subclass 172.3, for example.

II. Claims 8-10, drawn to plant cells containing a DNA construct comprising a gene encoding a desaturase or a thioesterase, classified in Class 435, Subclass 240.4, for example.

III. Claims 11-16, drawn to a DNA construct comprising a plant β-ketoacyl-ACP synthase gene and plant cells containing the construct, classified in Class 536, Subclass 23.6, for example.

IV. Claims 25-26, 36-37 and 39, drawn to a method of using a DNA construct comprising a plant β-ketoacyl-ACP synthase gene, classified in Class 435, Subclass 240.45, for example.

V. Claims 27-32 and 38, drawn to plant cells and plant seeds having modified fatty acid composition, classified in Class 800, Subclass 250, for example.

VI. Claims 40-41, drawn to oil, classified in Class 426, Subclass 601, for example.

Claims 10, 26, 37 and 39 are linking claims.

The inventions are distinct from each other because each involves a technical feature not required by any of the other inventions. The invention of Group I involves a bacterial gene not required by the other groups. The invention of Group II involves two other genes not required by the other groups. The inventions of Groups III and IV require a plant gene not required by the other groups. The invention of group IV involves processes for growing plant cells under conditions for transcription and for evaluating the change in fatty acid composition, wherein such processes are not required by the invention of Group III. The invention of Group V requires whole plant regeneration processes not required by any of the other groups. Furthermore, plants containing altered fatty acid profiles could be produced by processes other than those recited by the other groups, such as by conventional breeding procedures or by transformation with different genes. The invention of Group VI requires oil isolation techniques not required by any of the other groups. Furthermore, oil containing modified fatty acids could be obtained by processes other than those recited by the other groups, such as chemical modification of oil produced by wild type plants or isolation from modified plants produced by traditional breeding techniques. Accordingly, the claims are not so linked by a special technical feature within the meaning of PCT Rule 13.2 so as to form a single inventive concept.